Genetic Diversity in a Commercial Black Soldier Fly, *Hermetia illucens* (Diptera: Stratiomydiae), population

by

Rozane Badenhorst

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in the Faculty of AgriSciences at Stellenbosch University

Supervisor: Dr Clint Rhode

Co-supervisor: Dr Aletta E Bester-van der Merwe

March 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2017

Abstract

The applications of black soldier fly (BSF), *Hermetia illucens*, as a first-generation biotechnology address two global problems - the pressure on food security arising from unsustainable food production practices which exhausting fisheries resources by fishmeal production; and the vast accumulation of organic waste. Understanding the phenotypic and genetic changes experienced when establishing colonies of *H. illucens* is important for successfully establishing new colonies and managing already existing mass-rearing facilities. The aim of this study was to assess the changes in genetic diversity coupled with the phenotypic changes associated with colony establishment. Commercially-relevant phenotypic parameters were measured; including pupae weight, percentage eclosion, post-mating longevity, hatchability and egg clutch sizes over six successive generations (F0-F5). Hatchability, clutch size and pupae weight increased from F1 to F4, but a decline was apparent from F3 to F4. Post-mating longevity dramatically decreased from F1 (100%) to F4 (33%) and the F5 generation did not develop beyond the pre-pupae stage resulting in colony-collapse. The phenotypic changes can be explained by a combination of environmental, physiological and genetic effects. The positive trends observed during earlier generations (F0-F1) may reflect the population reacting to relaxed natural selection. While, at the end of the study period (F3-F5), natural selection in captivity resulted in a decrease in phenotypic variables. The colony collapse in the final generation, despite constant environmental conditions, possibly reflects the deleterious effects of random genetic drift and inbreeding depression**.** For the assessment of the genetic processes which may have resulted in the observed phenotypic changes, ten microsatellite markers were developed and characterised. Marker utility was tested on 37 wild individuals using three multiplex-PCR reactions. The number of alleles for each locus ranged from 4 to 21. Polymorphism information content ranged from 0.52 to 0.90, while observed and expected heterozygosity ranged from 0.30-0.65 and 0.55-0.91, respectively. Furthermore, the markers are useful in individual identification (P_{ID} = $1.2x10^{-11}$and P_{IDsib} = $1.1x10^{-4}$) and parentage analysis (P_1=

1; P_2 = 1). The markers proved useful in the assessment of genetic diversity and were used to understand the genetic mechanisms underlying phenotypic trends previously described. The assessment of genetic diversity revealed significant population differentiation as described by F_{ST} estimates across the experimental generations ($P < 0.01$), accompanied by a significant loss of genetic diversity. The final generation indicated a significantly higher number of private alleles and significantly lower observed heterozygosity in comparison to preceding generations ($P < 0.05$). Population bottlenecks and inbreeding were evident in the estimation of effective population size (N_e) and relatedness. Phenotypic parameters which significantly correlated (Pearson's *r*) to changes in genetic diversity were pupation (%), clutch size, pupae weight and female post-mating longevity ($P < 0.05$). Results indicate the effects of inbreeding depression and genetic drift were amplified by a decrease in effective population size and increased relatedness amongst individuals. In summation, the assessment of changes in phenotypic measurements and genetic diversity during the establishment of an *H. illucens* colony, gave valuable insight into the genetic processes which occur during colony establishment; and are applicable to colony establishment and the management of already existing colonies.

Opsomming

Die toepassings van *Hermetia illucens* as 'n eerste-generasie biotegnologie is daarop gemik om twee probleme op te los; die druk op voedselsekerheid wat voortspruit uit onvolhoubare voedselproduksiepraktyke (bv. die uitputting van visseryhulpbronne om vismeel te produseer) en die opeenhoping van organiese afval. Dit is van waardevolle belang om die verandering van kommersiële-relevante fenotipes en mate van genetiese diversiteit tydens kolonie-stigting in *H. illucens* te verstaan en toe te pas wanneer nuwe kolonies gestig word of om reeds bestaande kolonies te bestuur. Die doel van hierdie studie was om die fenotipiese en genetiese veranderinge tydens kolonie-stigting te bestudeer. Kommersiële-relevante fenotipiese mates soos aanwysers van kolonie-fiksheid, -vrugbaarheid en -gesondheid was gemeet, insluitende papie-gewig, langslewendheid, uitbroeivermoë van eiers en eier-broeiselgroottes. Hierdie eienskappe was gemeet oor ses opeenvolgende geslagte (F0-F5). Uitbroeivermoë, eierbroeiselgroottes en papiegewig het toegeneem vanaf F1 tot F4, maar 'n afname word waargeneem vanaf F3 tot F4 vir al drie veranderlikes. Die langslewendheid van wyfies het ook dramaties afgeneem van F1 (100%) tot F4 (33%) en die F5 geslag het nie verder as die voor-papiese stadium ontwikkel nie, wat gelei het tot die ineenstorting van die kolonie. Die tendense in fenotipiese veranderinge kan verklaar word deur 'n kombinasie van omgewings-, fisiologiese- en genetiese effekte. Die positiewe tendense tydens vroeëre geslagte (F0-F1) kan toegeskryf word aan geleidelike aanpassing tot die kunsmatige omgewing. Daarteen, kan die afname in fenotipiese veranderlikes tydens die einde van die studietydperk (F3-F5) veroorsaak word deur die kolonie-onderhoudstrategie. Die ineenstorting van die kolonie in die sesde generasie, ten spyte van voortdurende konstante omgewingstoestande, verwys na die nadelige gevolge van genetiese drywing en gevolglike intelingsdepressie. Die volgende doelwit was om gereedskap te ontwikkelvir die ondersoek van genetiese verandering tydens massa-grootmaak. 'n Paneel van tien mikrosatelliet merkers was ontwikkel en gekarakteriseer. Die toepassing van mikrosatelliet genotipering was getoets in 37 wilde individue

gdeur middel van drie multipleks-Polymerase Ketting Reaksie (PKR)-reaksies. Die aantal allele vir elke lokus het gewissel van 4 tot 21. Polimorfisme inligtinginhoud het gewissel van 0.52 tot 0.90, terwyl waargeneme en verwagte heterosigositeit gewissel het tussen 0.30-0.65 en 0.55-0.91, onderskeidelik. Die merkers was nuttig vir individuele identifikasie (P_{ID} = $1.2x10^{-11}$ en P_{IDsib} = $1.1x10^{-4}$) asook ouerskap-analise (P_1 = 1; P_2 = 1). Die merkers wat hier aangebied word, sal nuttig wees vir verskeie toepassings, insluitend die beoordeling van genetiese diversiteit tydens die stigting van *H. illucens* kolonies. Die merkers was in die huidige studie gebruik om die geniese meganismes wat fenotipiese tendense dryf, te ondersoek. Betekenisvolle populasie-differensiasie soos beskryf deur F_{ST} skattings oor die eksperimentele geslagte ($P < 0.01$) was waargeneem, tesame met 'n beduidende verlies van alleliese rykheid en heterosigositeit. Die finale geslag het ook 'n aansienlike groter aantal private allele en laer heterosigositeit getoon, in vergelyking met vorige geslagte ($P < 0.05$). Die berekening van effektiewe populasiegrootte (N_e) en ondersoek na verwantskap dui op ’n populasie bottelnek en inteling. Fenotipiese mates wat betekenisvolle korrelasier (Pearson se *r*) met veranderinge in genetiese diversiteit getoon het, was pupasie (%), eier-broeiselgrootte, papie-gewig en langslewendheid van wyfies na paring ($P < 0.05$). Die resultate dui aan dat die gevolge van intelingsdepressie en genetiese drywing versterk is deur 'n toenemende verwantskap tussen individue en klein effektiewe populasiegrootte. Ten slotte, die bevindinge van hierdie studie, sal nuttig wees in die toekomstige vestiging van *H. illucens* kolonies, asook vir die bestuur van reeds bestaande kolonies.

Acknowledgements

Firstly, I thank my supervisors, Dr Clint Rhode and Dr Aletta Bester-van der Merwe for their support, guidance and patience throughout this project. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. (Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF). Agriprotein Technologies (Pty) Ltd research and development facility in Cape Town, South Africa is thanked for the funding of this project and facilitation of colony establishment – thank you Dr Cameron Richards and Dr Ian Banks for being patient while I bugged you for all kinds of advice, and of course Ms Noxolo Mahashe for her contributions during colony maintenance. Dr Christine Picard and Anne A Andere from Indiana University-Purdue University (IUPUI) are thanked for their willingness to collaborate on the genome assemblies and development of microsatellite markers. Mr Peter Barnard from vermicomposters.com is also thanked for providing the founding members of the experimental colony. My gratitude also goes out to all the members of the Molecular Breeding and Biodiversity group; each member is appreciated for their support all the way from DNA extractions to genotyping and data analysis. To my family and friends, thank you for your support and believing in me at times when I did not know what I was doing or how I was going to do it; Mom, Dad, Franna, Ally, Martin, thank you!

Table of Contents

List of Figures

List of Tables

List of Abbreviations

%	Percentage
(Pty) Ltd	Property Limited
°C	Degree Celsius
µ	Micro-
µM	Micromolar
A	Adenine
bp	Base pair
BSF	Black soldier fly
C	Cytosine
CA	California
Ca-CO_3	Calcium carbonate
cm	Centimetre
CTAB	Cetyl trimethylammonium bromide
DLX6	Distal-less homeobox 6
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
e.g.	*exempli gratia* (for example)
EST	Expressed sequence tag
et al.	*et alia*
EtBr	Ethidium bromide
F	Filial generation
FAO	Food and Agriculture Organisation
FIASCO	Fast isolation by AFLP of sequence containing repeats
F_{IS}	Wright's Fixation Index (individual relative to sub-

	population, equal to inbreeding coefficient, *f*)
F_{IT}	Wright's Fixation Index (individual relative to total population)
F_{NULL}	Null allele frequency
F_{ST}	Wright's Fixation Index (sub-population relative to total population)
G	Guanine
g	Grams
H	Height
H_E	Expected heterozygosity
H_O	Observed heterozygosity
hrs	Hours
HWE	Hardy-Weinberg equilibrium
i.e.	*id est* (that is to say)
IUPUI	Indiana University- Purdue University, Indianapolis, USA
JH	Juvenile hormone
Kb	Kilo-base
kg	Kilogram
L	Length
L1-L6	Larval instar 1 - 6
M:F	Male to female ratio
MAS	Marker assisted selection
$MgCl_2$	Magnesium chloride
min	Minute
mL	Millilitre
mM	Milimolar

N	Number of (sample size)
N_A	Number of alleles
NCBI	National Center for Biotechnology Information
N_e	Effective population
NGS	Next generation sequencing
P	Probability value
PCR	Polymerase chain reaction
pH	Potential of hydrogen
PIC	Polymorphism information content
P_{ID}	Probability of identity
P_{IDsib}	Probability of sibling identity
PTTH	Prothoracic hormone
QTL	Quantitative trait locus
r	Pearson's correlation coefficient
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RH (%)	Percent relative humidity
R_S	Allelic richness
R_{SP}	Private allelic richness
s.d.	Standard deviation
SA	South Africa
sec	Second
SIT	Sterile insect technique
SRA	Sequence read archive
T	Thymine
T_a	Annealing temperature

Taq	*Thermus aquaticus* polymerase enzyme
TBE	Tris/Borate/EDTA
TD-PCR	Touchdown PCR
™	Trademark
U	Units (enzyme)
UK	United Kingdom
USA	United States of America
USD	United States Dollar
W	Width
w/v	Weight per Volume
ρ	Spearman's Rho correlation coefficient

"If I have seen further, it is by standing on the shoulders of giants."

-Isaac Newton, 1676

Chapter 1: Literature Review

With the anticipated continued growth of the global population, the global food demand is expected to increase for at least another 40 years. The ability to produce enough food to sustain population growth is affected by growing competition for land, water, and energy. Moreover, fisheries are overexploited to sustain agri- and aquacultural practices (FAO, 2014). In 2002, fishmeal and fish oil were primarily used for exhausive food production (24% for pigs, 22% for poultry and 46% for aquaculture) (Alder *et al.*, 2008). Moreover, the annual average global *per capita* consumption of fish increased from 9.9 kg to 17 kg between the years 1960 and 2000, and further expanded to 18.9 kg in 2010 (FAO, 2014). Meanwhile, the world population is expected to reach 9 billion by the year 2050, while fish production from capture fisheries is believed to be stagnant or declining (Rana *et al.*, 2009). Not only is the pressure on assuring food security increased, but so is the need to reduce excessive production of organic waste. It is estimated that as much as half of the food grown globally is lost or wasted both before and after it reaches the consumer. Most of this waste is fresh fruits and vegetables, as well as other perishables such as bakery and dairy products, meat and fish (Lundqvist *et al.*, 2008; Parfitt *et al.*, 2008). A multifaceted and collaborative global strategy is required to ensure sustainable and equitable food security, different components of which have been explored thoroughly (Parfitt *et al.*, 2008). Fortunately, there are many solutions to produce more food efficiently and sustainably, while efficiently managing waste accumulation. Exploiting the ways in which organic waste is naturally recycled into secondary protein products allows the efficient reduction of waste while producing animal protein. The applications of black soldier fly (BSF), *Hermetia illucens*, as a first - generation biotechnology are beneficial from two perspectives. Firstly, BSF larvae feed voraciously on organic waste such as animal manure and food waste. Secondly, larvae are efficient at bioconversion, converting their food-source into larval mass, which is high in protein and fat. These applications can be combined into one system in which waste is recycled to produce a sustainable animal protein source (Li *et al.*, 2011).

1.1. *Hermetia illucens*

Soldier flies (Stratiomyidae) are a family of the lower Brachycera suborder (higher flies), in the order Diptera. They are recognized mainly by their wing venation, with their radial veins concentrated in the anterior part of the wing and a small discal cell from which medial veins radiate (Figure 1.1) (Woodley, 2009). Although over 370 species occur in Africa, soldier flies have not been taxonomically revised in this region and the true diversity of species may be significantly higher (Villet, 2011). Among the Stratiomyidae, the genus *Hermetia* comprises 76 species in Nearctic, Neotropical, Afrotropical, Australasian and Oriental Regions (Üstüner *et al.*, 2003; Rohacek and Hora, 2013). *Hermetia illucens* is a species of focus in Africa as a result of its forensic and bio-remedial importance (Lord *et al.*, 1994; Villet, 2011). Global records of *H. illucens* indicate an increased frequency of encounters in Europe from 1950 – 1960 when more specimens have been sampled, but this is not a true indication of their abundance. It was first recorded in southern Europe in 1926. Preceding these recordings, the first record of BSF in South Africa was in 1915 (Picker *et al.*, 2004; Marshall *et al.*, 2015). The general consensus is that the spread of *H. illucens* was dependent on maritime transport that likely played a role in repeated, accidental introductions along coastlines and islands - colonising the KwaZulu-Natal province from South America (Picker *et al.*, 2004). The tendency of this species to colonise organic wastes leads to the assumption that fruit and vegetables on ships were invaded by BSF. Currently, molecular evidence supporting the biogeography of *H. illucens* and its colonisation of South Africa is lacking.

1.2. *Hermetia illucens* biology and life-history

The insect order, Diptera, forms part of the holometabolic insect groups, characterised by developing through four main stages which do not resemble the mature adult; namely, the egg, adult, larva and pupa (as opposed to hemimetabolic insects which go through nymphal instars, resembling the mature adult) (Figure 1.2).

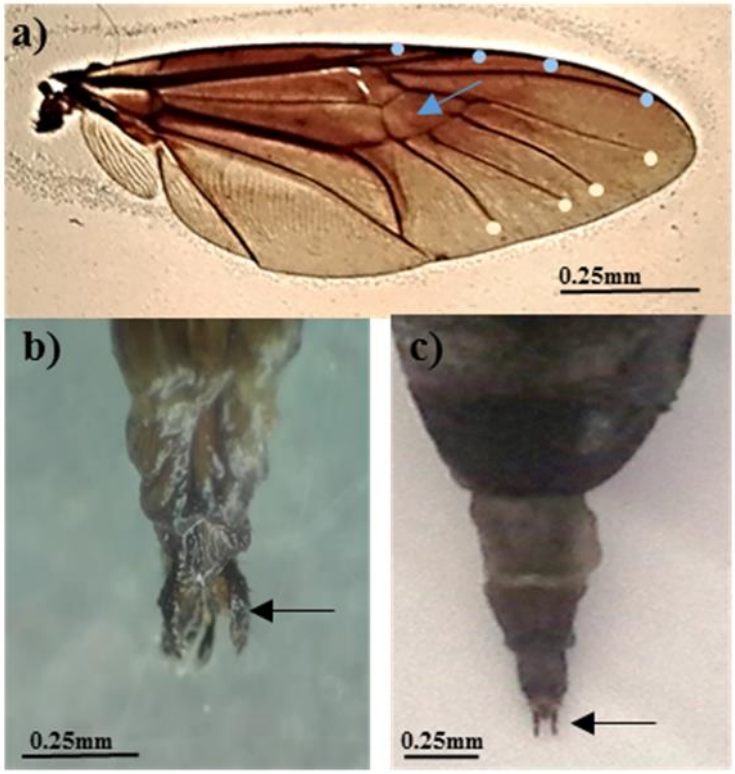

Figure 1.1: simple representation of the wing venation of *Hermetia illucens* (a). The radial veins are concentrated along the anterior margin of the wing (blue dots). The arrow indicates the discal cell, from which the medial veins radiate (yellow dots); b) male genitalia are identified by two pairs of posterior side lobes; c) female genitalia possess genital subtriangular furca at the base of the subgenital plate.

Sexual dimorphism in *H. illucens* is represented by genital structure (Figure 1.1). Male genitalia are short with two pairs of posterior side lobes, a pair of rims and a pair of very small gonostilos. Terminalia of a female consists of a long pair formed by two segments; the subgenital structure has a long plate in its distal portion with a pointed shape and genital subtriangular furca which are used to probe for appropriate oviposition sites (Oliveira *et al.*, 2016).

Adults rely on fat stores accumulated during the larval stage for dispersal, mating and ovarian development (Furman *et al.*, 1959; Tomberlin and Sheppard, 2002). Mating occurs two to four days post-eclosion and correlates to time of day (occurring mostly at dawn or dusk), while oviposition is also correlated to temperature and humidity. Oviposition typically occurs when temperatures exceed 26°C and only when relative humidity (% RH) exceeds 60 % (Tomberlin and Sheppard, 2002). Observations on their behaviour in the wild and in captivity have shown that adult males aggregate

in small numbers near secluded bushes (Copello, 1926). Lekking males exhibit territorial behaviour and engage in competitive displays where an individual grapples an approaching male, defending a small territory. The males will also attempt to mate with females visiting the lekking site (Tomberlin and Sheppard, 2001). Similar to *H. comstocki*, *H. illucens* females move away from lekking males once copulated to seek out oviposition sites (Alcock, 1990; Tomberlin and Sheppard, 2001).

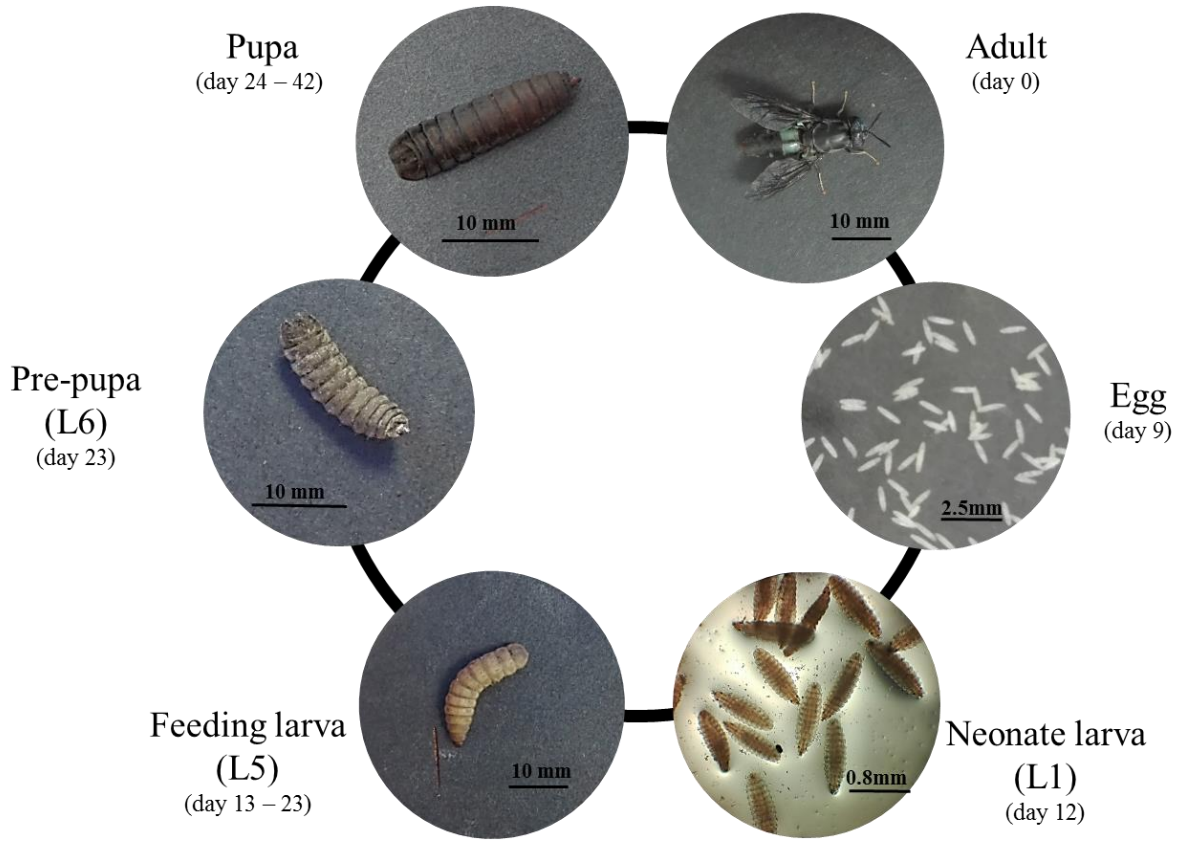

Figure 1.2: The duration of *Hermetia illucens* life-cycle is approximately 42 days (± four days) at 28°C. Adults are designated day 0 in the figure. After mating, eggs are oviposited in appropriate substrates (± day 9), after which neonate larvae hatch (± day 12). Larvae feed throughout six instars (L1 – L6) until reaching critical weight (± day 23), after which they cease feeding and seek pupation sites (± day 24). Pupariation begins and adults eclose on day 38 - 42.

Eggs are oviposited in clutches of 600 - 900 in crevices surrounding decomposing organic matter, which can range from animal carcasses to plant and faecal matter (Tomberlin *et al.*, 2002; Diener *et al.*, 2011; Banks *et al.*, 2014). Ovipostioning sites provide moisture which protects the eggs from desiccation and facilitates hatching, in addition to providing larvae with an immediate food source

after hatching (Holmes *et al.*, 2012). Eggs hatch within four days (102-105hrs at 27°C) and the neonate larvae begin feeding (Booth and Sheppard, 1984; Holmes *et al.*, 2012). The larvae feed for approximately two weeks, accumulating enough resources to reach critical weight, while developing through six larval instars (L1-L6). The critical weight is the minimal mass at which pupation can occur in a normal time-course (eight to 14 days in BSF). Once critical weight is reached, larval growth ceases as the secretion of ecdysteroids is stimulated by a sequence of endocrine and physiological events (Davidowitz *et al.*, 2003; Smykal *et al.*, 2014; Noriega, 2014). Juvenile hormone (JH) inhibits the secretion of prothoracic hormone (PTTH) and ecdysteroids. Once critical weight is reached, the level of JH decreases and metamorphosis begins (Nijhout and Williams, 1974). The decrease of JH titre results from the deactivation of the *corpora allata*, the glands which secrete JH. The subsequent secretion of PTTH then stimulates the secretion of ecdysteroids, which cause the larvae to cease feeding and commit to pupariation. The sixth instar is also referred to as "post-feeding larva" or "pre-pupae" in BSF. Aside from pre-pupae (L6), preceding instars are difficult to distinguish by eye in BSF. The size of the head, as well the distance between eyes and antenna can be used as an indicator of the larvae age and instars; this can be determined by microscopy and monitoring moulting events (Figure 1.3). For a detailed description on the external morphology of *Hermetia illucens* throughout development, see Oliveira *et al.* (2016).

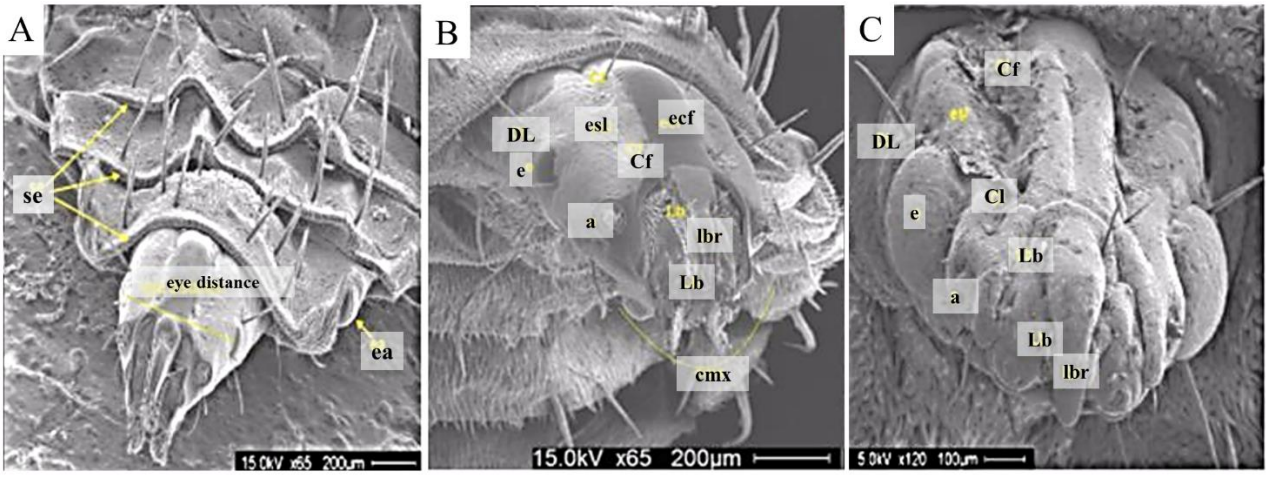

Figure 1.3: Head capsule and first thoracic segment of the larvae in the first and third instars. A) Distance between eyes, detail of cilia (se), anterior spiracles (ea). B). First instar. Detail of seta frontoclypeal (Cf), lateral sclerite (esl), frontoclypeal sclerite (ecf), seta dorsolateral (DL), compound eye (e), antenna (a), seta labrum (Lb), labrum (lbr), maxilla and mandible complex (cmx). C). Head of larva in the third instar. Detail of seta frontoclypeal (Cf), seta dorsolateral (DL), compound eye (e), antenna (a), seta labrum (Lb), labrum (lbr). Images have been adapted from Oliveira *et al.* (2016).

Following the cascade of endocrine events described, L6 larvae cease feeding and migrate from the feeding larval aggregation to suitable pupation sites. In the wild, this would typically be in the drier areas of soil, away from the food source. The availability of a pupation substrate is vital; without it, the pre-pupae will not begin pupariation. Pupariation is the period between the time that the larvae cease feeding and complete immobilization and reduction in length of the larvae takes place. A reduction in larval mobility and a retraction of the body segments gradually occurs. The cuticle becomes progressively more opaque, pigmented and sclerotized (Barros-Cordeiro *et al.,* 2014). The pupa is formed within the last larval skin, which is used as a hard cocoon (puparia) impregnated with plates or calcium carbonate crystals ($CaCO_3$) (Woodley, 1989). The suitability of pupation substrates depends on the humidity and compactness of the substrate. Moist substrates with lower compaction density are preferred, for example wood shavings or vermiculite (Holmes *et al.,* 2013). Pupation or intra-pupal development includes events that occur from larval-pupal apolysis until the emergence of the adult fly. The process of pupation occurs in the first six hours, extroversion of the head and thoracic appendages then takes place between hours nine and 21, and the pharate (mature

adult) develops 21 hours after completing pupation (Barros-Cordeiro *et al.,* 2014). Pupae eclose as adults eight to 14 days post-pupariation, after which the adults fully sclerotize and mature over the course of 48 hours and the cycle begins again (Tomberlin *et al.,* 2002; Barros-Cordeiro *et al.,* 2014).

1.3. *Hermetia illucens* in biotechnology

Early accounts of the black soldier fly have described it as a sanitary pest with little economic importance (Bradely, 1930). However, interest in the biology and life-history of BSF has since grown with the realisation of the fly's bio-remedial potential, converting organic wastes into a sustainable protein source comparable to that of conventional fishmeal (Figure 1.4). *Hermetia illucens* is capable of converting poultry manure to a 42% protein, 35% fat feedstuff (Sheppard *et al.,* 1994). Varying waste-to-biomass conversions have been reported, ranging from 8-22.9%, depending on the waste source (Banks *et al.,* 2014; Lalander *et al.,* 2014; Diener *et al.,* 2011; Sheppard *et al.,* 1994). Organic waste sources which have been experimentally tested range from palm kernel meal, pig liver, livestock manure, kitchen waste and rendered fish to human faecal waste (Hem *et al.,* 2008; Lalander *et al.,* 2013; Choi *et al.,* 2012; Diener *et al.,* 2011; Kalová and Borkovcová, 2013; Nguyen *et al.,* 2013; Zhou *et al.,* 2013; Banks *et al.,* 2014; Nguyen *et al.,* 2015).

By-products of BSF-mediated waste reduction in the form of larval fat or protein can then be used as biofuel and animal feed supplements (Figure 1.4) (Bondari and Sheppard, 1981; Li *et al.,* 2011; Zheng *et al.,* 2011; Zheng *et al.,* 2012). The potential of milled or whole dried larvae as an animal feed supplement has been demonstrated as early as 1973 (Hale, 1973). In fact, larvae have been fed experimentally to several animal species, including poultry, the mountain chicken frog (*Leptodactylus fallax*), American alligator (*Alligator mississippiensis*), channel catfish (*Ictalurus punctatus*), blue tilapia (*Oreochromis aureus*), juvenile turbot (*Psetta maxima*) and rainbow trout (*Oncorhynchus mykiss*), with larvae used to supplement soybean or fish meal in formulated diets

(Hale, 1973; Bondari and Sheppard, 1987; Sheppard and Newton, 1994; Bodri and Cole, 2007; St-Hilaire *et al.*, 2007; Dierenfeld and King, 2008; Kroeckel *et al.*, 2012; Kenis *et al.*, 2014).

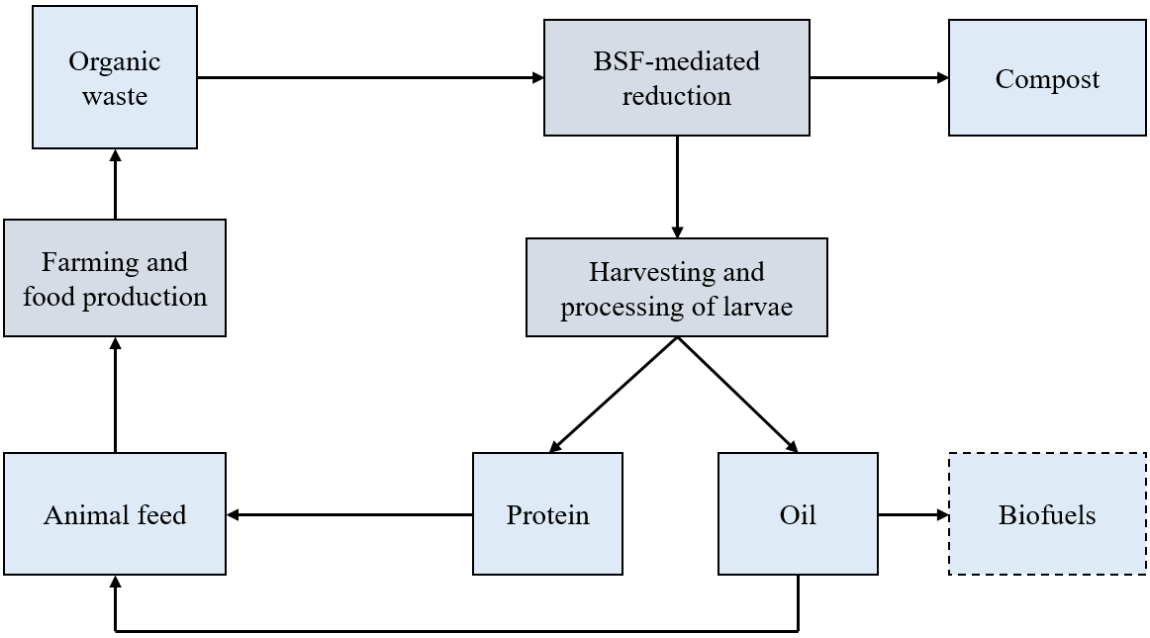

Figure 1.4: An outline of the process of biotechnological application of BSF. Organic waste is reduced and recycled by feeding it to the larvae. Larvae convert the nutrients from the waste into protein, while acting as composting agents. Processing involves separating protein and fats which can be used as biofuels or as additives in animal feed. The feed is processed through animal farming as animal feed (e.g. poultry and fish). The by-products of farming and food production (such as offal), are again fed back into the BSF-recycling system.

Hermetia illucens is an especially attractive option for commercial mass-rearing because of attributes which make it easier to comply with sanitary regulations. For example, the adult fly does not come into contact with any degrading or fresh organic material. Therefore, unlike other species considered for this technology (*i.e.* house fly (*Musca domestica*), blue bottle (*Calliphora vomitoria*) and other blow flies (*Chrysomya spp.*), *H. illucens* is not considered unsanitary or a mechanical vector of disease (Schremmer, 1986; Leclercq, 1997; Charlton *et al.*, 2015). In fact, larval extracts have been found to be effective in the prevention of *Klebsiella pneumonia*, *Neisseria gonorrhoeae* and *Shigella sonnei* growth (Choi *et al.*, 2012; Chu *et al.*, 2014). The aforementioned characteristics make BSF an ideal candidate for commercial mass-rearing. This has already been demonstrated

globally in various BSF rearing facilities, where standard insect mass-rearing strategies are implemented (Marshall *et al.*, 2015).

1.4. Insect mass-rearing

In comparison to conventional livestock, insects have a higher feed-conversion efficiency, meaning they require less feed-input to produce 1kg of biomass (Nakagaki and Defoliart, 1991).They are also highly fecund, mostly omnivorous, and as nutritious as conventional livestock, while occupying less space during the rearing process (Rumpold and Schlüter, 2013). Candidate insects are considered based on their size, social behaviour, safety, epidemic tendencies, reproductive and adaptive potential, nutritional benefits, potential for storage, and marketability (Schabel, 2010).

Once an appropriate candidate is selected, it is essential to develop appropriate safety and quality monitoring for the automation of insect (protein) production, to limit production costs and ensure product safety. The first section in Figure A.1 (Appendix A) highlights the production of animal protein by the process described in Figure 1.3. Regardless of the life-stage targeted for the end product, holometabolous insects are mass-produced following the same rationale. The first attempts to truly upscale mass-rearing of insects followed Knipling's (1955) proposal to implement sterile insect technique (SIT). Appropriately, as interest grew in the development of insect mass-rearing programs, production efficiency was clearly defined. To achieve optimal production efficiency in insect mass-rearing programs, the goal is to produce the maximal number of insects while expending minimal man-hours and space, in as little time as inexpensively possible (Finney and Fisher, 1964). In order to be classified an insect mass-production system, mass-rearing facilities must reach the daily production of one tonne of product as decided on the FAO technical consultation meeting in 2012 (Vantomme *et al.*, 2012).

To achieve a minimum daily production turnover, each checkpoint along the entire production chain must be optimised and monitored. Optimality and efficiency is monitored by implementing quality

control protocols. High egg production and hatchability, shortened larval development, synchronized pupation, high larval or pupae weight, a high bioconversion rate, low feed costs, high tolerance for diseases, ability to live in high densities, and a high quality protein are aimed for. Additionally, it is pertinent that rearing conditions are controlled throughout the rearing process along the production chain (Figure A.1) (Appendix A). Environmental factors influencing the rearing process include temperature, light/illumination, and humidity, ventilation, rearing container properties, larvae/population density, oviposition media, food and water availability, food composition, food quality as well as microbial contamination (Peters and Barbosa, 1977; Scriber and Slansky, 1981; Singh, 1982; Sharaby *et al.,* 2010; Tchuinkam *et al.,* 2011; Vantomme *et al.,* 2012; Oonincx *et al.,* 2016).

Hermetia illucens is widely mass-reared for a variety of commercial uses in the United States, Costa Rica, Europe, and South Africa (Marshall *et al.,* 2015). The basic process of mass-rearing *H. illucens* is similar to that of other holometabolic insects (Figure A.1) (Appendix A). Eggs are collected and weighed from cages in which adults mate. A predetermined weight of eggs is distributed into initial feedings bins, controlling the larval feeding-ration- which has been optimised to facilitate reaching critical weight with minimal uneaten organic waste (Diener *et al.,* 2009). Eggs are then randomly subdivided into batches which will carry on to become product, while a fraction is set aside to re-establish adult colonies. After hatching, neonate larvae are nursed on specialised diets until they reach a more robust stage of development (~L2-L3). Larvae are then moved to one large collective trough and fed a range of organic waste. Offspring from different cages and collection days is combined during this phase. Pre-pupae are collected and allowed to pupate. Cages are then again stocked with equal portions of pupae, to maintain a predetermined optimal density of adults per cage. This is similar to pupae raffling described by Gilchrist *et al.* (2012) in SIT facilities. The adult density refers to catering for the maximum number of adults in a cage to increase egg production, while providing enough space for lekking and mating to occur.

1.5. Population genetics and domestication in insect mass-rearing

Domestication is defined by Price (1984) as "the process by which a population of animals becomes adapted to man and to the captive environment by some combination of genetic changes occurring over generations and environmentally induced developmental events recurring during each generation". Population genetics theory predicts the extent of adaptation to captivity (and subsequent domestication) depends upon, effective population size, genetic diversity and selection intensities (Frankham, 2008).

When a population is isolated from the wild in an artificial environment, it experiences a bottleneck. The decrease in effective population (N_e) size during the founding event causes significant declines in genetic diversity by the founder effect (Parker, 2005). The loss of genetic diversity has been quantified on multiple occasions during colony establishment by observing an initial loss of alleles, changes in allele frequencies, decrease in allelic richness and decreased heterozygosity in studies of *Anopheles gambiae, Bactrocera olaea* and *B. tryoni* (Norris *et al.,* 2001; Baeshen *et al.,* 2014; Gilchrist *et al.,* 2012). Castañé *et al.* (2014) previously demonstrated the importance of sufficient founder population size when establishing captive colonies of the predatory biocontrol agent, *Orius laevigatus* (Hemiptera: Anthocoridae), for example. In their study, populations with a significantly larger number of founding couples (50 and 10 couples, compared to one) were able to persist in captive rearing conditions, while the smaller founding population (one couple) deteriorated. Moreover, origin of the founding population is equally important to consider, to ensure a large genetic diversity of the founding population by sampling from geographically segregated populations; as demonstrated in *O. laevigatus* and *Diabrotica virgifera virgifera* (Castañé *et al.,* 2014; Li *et al.,* 2014).

The size (N_c) and genetic diversity of the founding population influences the degree to which genetic drift and inbreeding occur. The decrease in N_e associated with the founding event increases

the chances of losing rare alleles by random genetic drift which further decreases heterozygosity. Small populations are also prone to inbreeding, resulting in a further decrease in heterozygosity.

The population also experiences multiple selective sweeps throughout colony establishment. Heterozygosity is decreased by natural selection in captivity. Zygouridis *et al.* (2014) observed a 50% loss heterozygosity by F11 in a colony of *B. olaea*. This loss of genetic diversity, reflects the loss of the "wild" character of the strain and possibly of its natural vigour. In the same study, population size increased significantly in later generations, until reaching a stable population size. Increasing the genetic diversity of the base population, allows for increased capacity to adapt and more tangibility during downstream artificial selection.

While natural selection in captivity results in bottlenecks, relaxed natural selection refers to the alleviation of natural selection pressures and applies to traits which would be beneficial in the wild, but are not apparently beneficial in captivity. An example of this in insect domestication is the loss of natural selection pressure brought on by starvation risks, when formulated feeds are provided *ad libitum*. This is a by-product of artificial rearing, as individuals with higher feeding efficiencies produce larger biomass in a shorter time-frame (Ekesi, 2007).

The third type of selection to be considered when studying domestication is artificial selection. Artificial selection consists of the selection of breeding individuals with favourable traits and results in different breeds or strains. Unintentional artificial selection occurs during the “pupal raffle” process described by Gilchrist *et al.* (2012). Genetic diversity is lost during inadvertent selection for increased pupae size during pupal raffling. The dependence of this strategy in mass rearing-systems to sustain breeder colonies, results in the continual selection for increased fecundity throughout the program, as pupae weights have previously been correlated to egg production (Pastor *et al.*, 2011). Moreover, simulations have shown the initial skewed fecundity-distribution caused by this process (Gilchrist *et al.*, 2012), and that this may be explained by the resource allocation theory. When

predicting the changes resulting from domestication, understanding the resources allocation theory is also helpful. The theory maintains that under selection within a particular environment, the resources used by the animal are distributed between the important traits for breeding and production within that environment (Beilharz *et al.*, 1993). Additional selection pressures which increase performance of production traits typical of conventional livestock farming (e.g. milk production) would result in a shift of resource allocation from reproductive traits. An initial loss of natural vigour occurs during relaxed natural selection. Then, by artificial selection, the ability to reproduce in captivity increases, and the allocation of resources shifts to production performance. In conventional livestock, this becomes problematic, as some of the high-producing individuals have difficulties in reproduction as the shift in resource allocation occurs (e.g. dairy cows which have reproduction problems - Rauw *et al.*, 1998).

With appropriate artificial selection strategies, phenotypic changes beneficial to mass rearing can be achieved. However, careful consideration must be given to selection strategy, as both intentional and unintentional forms of selection can result in a decrease in overall heterozygosity. The reduction in effective population size caused by selecting for individuals adapted to the artificial rearing environment, further increases the probability of inbreeding. Inbreeding depression (a decrease in fitness as heterozygosity decreases) then occurs (Nei *et al.*, 1975). Moreover, genetic drift would increase the expression of deleterious alleles, but the increase in expression of recessive alleles may also expose rare alleles to selection, potentially uncovering variation that facilitates adaptation to new environments (Knowles *et al.*, 1999; Reed *et al.*, 2003). In captive populations, inbreeding depression and random genetic drift can occur when the effective number of breeding individuals, or the effective population size, N_e, is mismanaged (Hill, 1982). On the other hand, artificial selection pressures have valuable effects when N_e is large enough for selection to overcome genetic drift (Crow and Kimura, 1970).

In insect mass-rearing, fitness-related traits are often also production-related traits. Pupae weight and adult size shows a general correlation to average clutch size in many insects. However, hatchability of the eggs may vary despite clutch size, this has previously been demonstrated in *Hermetia illucens* (Honěk, 1993; Tomberlin and Sheppard, 2002; Tomberlin *et al.*, 2002). Strictly controlling the environmental conditions conducive to hatchability, and selecting for individuals producing large clutch sizes, may be a preliminary strategy- while the relationship between hatchability and clutch size is further clarified.

1.6. Phenotypic changes associated with insect domestication

As a result of the genetic processes described, domestication has resulted in modifications of many traits determining the capacity of adaptation of animals, including behaviour, physiology and morphology. The changes discussed are relevant to mass-rearing as they either increase production outputs by increased fertility and fecundity or by increased overall biomass.

Shifts in feeding behaviour occurred within five generations of rearing *Bacterocera invadens* as a result of increased natural selection for utilising artificial feed. The increased feeding-efficiency results in an increased biomass production (Ekesi, 2007). Changes in mating behaviour which have been reported, include a change in time of mating in the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae). The time of wing vibration (mate signalling) and mating occurs later in wild than in mass-reared individuals. Additional changes in behavioural traits included flight ability, and dispersal distances which were higher in wild individuals. The combinations of these behavioural modifications indicate a shift toward earlier reproduction, while losing the ability to disperse, as it is not required in the artificial environment (Miyatake, 1998). Changes in male lekking behaviour have also been reported in the Queensland fruit fly, *Bactrocera tryoni*. Mass-reared males engaged in competitive displays with other males more frequently than wild males, and began mate-calling significantly earlier (Weldon, 2005). This further indicates the shift to increased frequency of

reproduction, but may result in decreased mate-preference. In other words, the frequency of sibling-mating may increase leading to decreased genetic diversity and decreased fitness. Additional life history changes have also been observed; Meats *et al.* (2004) found that by the fourth generation, *B. tyroni* reached sexual maturity and oviposited earlier than wild populations. However, egg production did not increase and was significantly lower than older captive strains. *Bactrocera cucurbitae* also showed heritable decreases in developmental time within 10 generations (Miyatake and Yamagishi, 1999) and, viability and oviposition patterns in *Ceratitis capitata* both increased significantly within the first 10 generations of domestication (Souza *et al.*, 1988).

Overall, shortening in the developmental and pre-oviposition periods, age of peak fecundity, ovarian development time, post-ovipositional life span, longevity, pre-mating period and re-mating intervals occurs over successive generations as selection for higher productivity takes place (which in the case of insect mass-rearing, may also be reproductive traits). As such, changes in reproduction have included an increase in hatchability, fecundity and % oviposition in a variety of species. Moreover, shortening of developmental periods has resulted in larger overall biomass as the insects become more efficient in feed metabolism - this can be viewed as relaxed natural selection while populations are fed optimal diets, while artificial selection has contributed to the shortening of developmental time (Souza *et al.*, 1988; Miyatake, 1998; Weldon, 2005; Ekesi, 2007).

Eventually, a level of domestication is achieved as the resulting newly colonized strain progressively adapts to artificial conditions and is selected to have traits conducive to mass-rearing. Colonised strains of *Anopheles* species, for example, are so extensively adapted to laboratory conditions that they fail to establish in the wild. Additionally, in the many Tephritidae examples described (*Bacterocera* and *Ceratitis spp.*), many traits respond to selection during the domestication process. Reasonably it is suspected that many genes are also under selection or that there are many traits controlled by the expression of single gene. Investigating the epistatic and pleiotropic interactions would further aid in the development of selection programmes (Flint and

Mackay, 2009). Achieving a balance between the genetic processes can only be achieved by expanding investigations using an array tools available to study genetic diversity. The impact of genetic management of captive colonies is improved by studying the population genetics throughout colony establishment. Understanding the levels of genetic diversity and population demographics improves the design of strategies to manage captive populations. There are a variety of molecular markers and population genetic estimates which assist in these designing these strategies.

1.7. Tools for monitoring genetic diversity

Originally, genetic diversity and differentiation were assessed by morphological traits, followed by allozymes (enzyme products of genes) and then at the DNA level, restriction fragment length polymorphism (RFLP). As PCR techniques became more prevalent and thereby more cost-effective, morphological and allozymatic approaches were outmatched (Gaudeul *et al.*, 2004). There are two main categories of PCR-based molecular markers: Type I markers are associated with genes of known function, while Type II markers are associated with unknown genomic sequences (Liu and Cordes, 2004). Type II genetic markers provide information in one of two ways: through a binary molecular phenotype represented by the presence/absence of given fragments (dominant) or by means of a genotype wherein both alleles of a diploid individual are revealed (codominant) (Anne, 2006). Microsatellites can be both Type I and Type II markers and have been previously developed for insect identification and population and evolutionary genetic studies. Microsatellite markers are particularly attractive, as they can be developed without prior knowledge of gene functionality and can be pooled in high-throughput multiplex PCR reactions.

1.8. Microsatellite markers

As described by Chambers and MacAvoy (2000), microsatellites are simple sequence repeat loci, consisting of two to six nucleotides repeated in tandem, in tracts of <1kb. They represent

codominant genetic loci; meaning that heterozygous genotypes are distinguishable from homozygous genotypes (Chistiakov *et al.*, 2005; Webster and Reichart, 2005). Microsatellites are ubiquitous throughout the genome of almost all living organisms, making the data acquired through using them comparable between studies (Chambers and MacAvoy *et al.*, 2000). Lagercrantz *et al.* (1993) and Ross *et al.* (2003) found that the frequency, mutation rate and motif of these repeat sequences varies considerably across genomes of different species; with the recent consensus being that their distribution might not be random. They are also longer and more common in vertebrates than invertebrates (Amos, 1999). Loci consisting of the dinucleotide repeat $(CA)_n$, are the most abundant in most genomes, with high levels of heterozygosity. Trinucleotide and tetranucleotide motifs occur less frequently and have lower levels of heterozygosity (Mueller and Wolfenbarger, 1999; Chambers and MacAvoy *et al.*, 2000). Kruglyak *et al.* (1998) found that an inverse relationship exists between relative abundance and repeat motif; likely as a result of higher slippage-rates experienced by shorter motifs (mono- and dinucleotides) when compared to longer ones. According to Primmer *et al.* (1997), there appears to be a disposition for the density and length of microsatellites to increase as genome size increases whereas. Deka *et al.* (1999), on the other hand, suggest that the genomic frequency of microsatellites is not directly related to genome size but rather a result of species-specific recombination rates and/or mutation or repair rates acting within the species under investigation. In insects, however, it has been stated that microsatellite length and frequency correlates with genome size (Hancock, 1996).

Patterns of microsatellite distribution across the genome appear to be similar in insects such as *Drosophila spp.* and *Anopheles gambiae* (Zheng *et al.*, 1993). With the exception of rare trinucleotide repeats occurring in coding regions, the distribution of microsatellites across intergenic areas is uniform (Hancock, 1999). Since coding exons typically lack microsatellites, or contain mostly trinucleotide repeats, the notion is that microsatellite variation is either deleterious or restricted to non-functional intergenic DNA. Recently, however, many reports have demonstrated

that a large number of microsatellites are located in transcribed regions of genomes, including protein-coding genes and expressed sequence tags (ESTs). One example of a phenotypic effect of microsatellites located within coding regions in insects, is the resultant of cell death cause by triplet expansion of a $(CAG)_n$ repeat in the Homeobox gene *DLX6* (Ferro *et al.*, 2001).

The mutation rates observed for microsatellite loci can be as high as 1.4×10^{-2} per generation (Talbot *et al.*, 1995; Wang, 2004). There are two likely mechanisms proposed for microsatellite mutation. The first proposed by Schlötterer and Tautz (1992) suggests that microsatellites evolve through frameshift mutations resulting from slippage events during DNA replication or repair; while Richards and Sutherland (1994) present a scenario of unequal recombination, this refers to both Mendelian recombination and transposition. Recombination-like processes that involve unequal crossover or gene conversion introduce mutations in the larger minisatellite sequences, but evidence that recombination would also contribute to microsatellite mutations is lacking (Berg *et al.*, 2003). Genomic microsatellite distributions are associated with sites of recombination (Majewski and Ott, 2000), likely as a consequence of repetitive sequences being involved in recombination rather than being the cause (Treco and Arnheim, 1986). Most tests for a correlation between recombination rate and microsatellite density or mutability have failed to confirm such an effect (Bachtrog *et al.*, 1999). Hence, the predominant mutation mechanism in microsatellite tracts appears to be "slipped-strand mispairing" and is thoroughly described by Eisen (1999). Depending on whether the newly synthesised DNA strand or the template DNA strand loops out, slipped-strand mispairing would either cause the gain or loss of repeat units. The inclination for either strand to loop out is partly dependant on the array sequence. Lai and Sun (2003) also found that when PCR slippage occurs, the probability for contraction of the repeat was higher than the probability for the repeat to expand and that a threshold repeat number exists under which the mutation rate becomes too small to be observed. The assumption from their observations in these two studies was that each repeat unit within a microsatellite mutates independently. Overall, observations reviewed by

Chambers and MacAvoy (2000) indicate an upward bias of mutation, in which case the gain of unit mutations occurs more frequently (or is repaired less frequently) than the loss of unit mutations. Overall, it seems there are still discrepancies surrounding various aspects of microsatellite biology and evolution such as their mechanism of mutation or their functionality in the genome.

1.9. Microsatellite isolation, amplification and analysis

Previously, microsatellites were isolated from genomic libraries; assembled by screening thousands of clones through colony hybridization with repeat-containing probes. This is a tedious approach and becomes problematic in taxa with low microsatellite frequencies, such as birds or plants (Cooper *et al.*, 1995; Di Maio and De Castro, 2013). Alternative strategies have been implemented to avoid this process of library construction. For example, randomly amplified polymorphic DNA (RAPD) approaches have been employed by either using repeat-anchored random primers or RAPD primers in combination with Southern hybridization of PCR bands with microsatellite probes. Another popular method was the fast isolation by amplified fragment length polymorphism of sequences containing repeats (FIASCO) (Zane *et al.*, 2002). In summation, the basic approach to isolating microsatellite loci previously involved digestion, hybridisation, cloning and sequencing with a multitude of strategies (Zane *et al.*, 2002). Rapidly expanding sequencing technology has alleviated the drawbacks associated with having to isolate microsatellites *de novo* for previously unstudied species; next generation sequencing (NGS), pyrosequencing (454 Life Sciences, Roche Diagnostics, Indianapolis, Inidians, USA), Illumina technology (Illumina, San Diego, California, USA) and pH-change sequencing such as Ion Proton technology (Life Technologies, Paisley, Renfrewshire, UK) have all progressively improved the efficiency and cost of isolating microsatellite loci (Di Maio and De Castro, 2013).

Once isolated, microsatellite loci are targeted by designing primers specific to flanking regions on either side of the repeat element. Primers are generally 18-24 base-pairs in length and highly

conserved in the genome. Polymerase chain reaction (PCR) is a practical method to target and amplify these regions. Separation of PCR products by size then allows for the assessment of variation in different populations. Allelic inheritance can be tracked form progenitor to progeny and in doing so, inter- or intraspecific population genetic data can be generated. The resolving power increases with number of markers used (Wang *et al.,* 2009). Electrophoresis is used to separate products by size. Agarose or polyacrylamide gels are able to detect polymorphism at different levels; however, resolution is lost when alleles differ by small increments, undetectable by gel electrophoresis. Capillary electrophoresis solves this problem and simultaneously, allows for the compilation of different markers into one multiplex reaction by fluorescent labelling. Once successful marker amplification and separation is achieved, the size of the alleles can be scored. There are many automated platforms available for allele scoring (genotyping). Platforms differ in throughput rate and the choice of which to use is therefore dictated by the magnitude of the project and available funding. Comparison of different techniques in microsatellite isolation and analysis are discussed in detail by Wang *et al.* (2009).

1.10. Applications of microsatellites

Microsatellite markers, although relatively costly to develop, are able to address a wide range of biological problems because of the high-resolution data they can provide. The cost of development is outweighed by the information that can be obtained. The size differences between alleles as a result of a varying number of repeat units at a given microsatellite locus, is the basis for this molecular marker polymorphism (Liu and Cordes, 2004). Owing to the characteristics of microsatellites described previously, a wide range of applications exist. Surveying genetic diversity of wild founding populations ensures the maximum genetic diversity can be included during the initial stages of colony establishment. Estimates of relatedness and parentage analysis also aids in the understanding of the organism's life-history which improves the management strategies of populations. Other than maintaining genetic diversity, production performance can also be

improved. A prime example includes marker assisted selection (MAS) by use of quantitative trait loci (QTL's). This approach involves genome-wide screening for loci associated with particular traits of interest. Using this approach has proven invaluable in selective breeding for a variety of taxa, including insects (Nagaraju, 2002; Collard *et al.*, 2005).

1.11. Short-comings in microsatellite analysis

The ability to process large sample sizes via multiplex PCR lends itself to data collection errors. Genotyping errors may occur frequently even under optimized PCR conditions and with high quality template DNA. This is especially true for data where repeat typing is limited or not feasible owing to cost or time constraints or limited DNA (Wang *et al.*, 2004). Microsatellite genotyping errors are common in large data sets; and may weaken downstream analyses- emphasising the need to correct for these errors within analysis techniques and software programs (Hoffman and Amos, 2005). Here the potential errors arising during genotyping are discussed. Owing to slipped-strand mispairing being the underlying mechanisms of microsatellite mutation, one of the complications occurring during the analysis of microsatellite data is the appearance of stutter patterns or stutter profiles. Secondary or shadow bands result in extended banding patterns (Miller and Yuan 1997; Mueller and Wolfenbarger, 1999). The problem of stutter profiles is observed to be most acute where dinucleotide repeats are concerned while greatly reduced in tri and tetra-nucleotide repeats; with further evidence reporting that longer microsatellite repeats are more prone to producing stutter than shorter microsatellites (Miller and Yuan, 1997; Mueller and Wolfenbarger, 1999). Stutter bands can be reduced and sometimes eliminated through the optimization of individual PCR conditions and through the use of "hot start" PCR (Mueller and Wolfenbarger, 1999).

Another frequent problem when using microsatellite loci is the occurrence of null alleles. Microsatellite null alleles are any allele at a microsatellite locus that repeatedly fail to amplify to detectable levels via PCR (Dakin and Avise 2004). In addition, the preferential amplification of

allele size variants due to the competitive nature of PCR results in more efficient amplification of shorter alleles. The larger allele is reported to "drop-out" and a null allele resulting from this preferential amplification is sometimes referred to as a "partial null". PCR slippage is a common hindrance to assigning alleles correctly in microsatellite genotyping, as they may overlap in heterozygous alleles. This forms a biased compound pattern leading to possible misinterpretation of the data as a homozygous pattern (Mueller and Wolfenbarger, 1999). A separate cause of null alleles involves inconsistent PCR results owing to poor DNA quality or quantity. The unpredictability of PCR amplification due to poor DNA quality makes it difficult to confirm if this is indeed the cause of the null allele (Dakin and Avise 2004). Since the presence of null alleles biases estimations of allele and genotype frequencies, it is important to accurately identify whether they are the source of homozygous excess and estimate their occurrence (van Oosterhout *et al.*, 2006). Chakraborty *et al.* (1992) explain that it is important to recognise that when populations are out of Hardy-Weinberg equilibrium, heterozygote deficiency will be detected. This may not always be as a result of null alleles but could also be due to biological factors such as population substructure, inbreeding or selection occurring at or near the locus. Null alleles are common in insects and should be treated with caution when carrying out population studies (Thoren *et al.*, 1995). For example, when sampling the founding members for a new facility, there may be an occurrence of null alleles in the data, it is important to ascertain whether this is as a result of perhaps homozygous excess or resultant of the other confounding effects mentioned above.

Importantly, the utility of microsatellites is limited by the application of various statistical methods concerned with analysing the data they produce (Chambers and McAvoy, 2000). Despite the potential shortcomings of microsatellite markers, they have been widely used to study a variety of the genetic mechanisms relevant to domestication and insect mass-rearing.

1.12. Rationale, Aims and Objectives

As a fast-developing industry, the recycling technologies employing *H. illucens* are rapidly expanding. Understanding the change of commercially-relevant phenotypes experienced when establishing commercial colonies of *H. illucens* could aid in the management and optimising of a mass-rearing system. More so, understanding the decline in genetic diversity associated with this process is invaluable. This study focussed on the temporal genetic changes occurring across generations during early attempts at domestication. The "longitudinal approach" described by Price (2002) was implemented. Estimates of genetic diversity could further improve effective management and breeding of captive populations for long-term, sustainable production.

The first objective of this study was to investigate the various phenotypic changes experienced when establishing a colony from wild individuals (Chapter 2). Phenotypic parameters relevant to commercial mass-rearing were measured; indicators of fitness, fecundity and colony health such as pupae weight, percentage eclosion, post-mating longevity, hatchability of eggs and egg clutch sizes were recorded.

The next objective was to develop tools to aid in the assessment of genetic change during domestication and mass-rearing BSF; and to supplement the findings found in Chapter 2 (Chapter 3). A panel of polymorphic microsatellite markers was developed and characterised to aid in the assessment of genetic diversity across experimental generations. These markers will represent a novel panel of microsatellite markers for *H. illucens*.

The final objective was to apply the microsatellite marker panel developed in Chapter 3 to understand the trends observed in Chapter 2. Aside from gaining an understanding of the phenotypic trends, Chapter 4 focused on trends in genetic diversity and gives insight in to population demographics across generations.

Chapter 2: The phenotypic changes associated with domestication of the black soldier fly, *Hermetia illucens* (Diptera: Stratiomyidae)

Abstract

During the process of domesticating insects for mass-rearing, a variety of morphological, behavioural and physiological changes occur. Assessing how commercially-relevant phenotypes change over successive generations under commercial production aids in the successful establishment of new colonies, as well as the management of currently existing colonies. This chapter aimed at investigating the various phenotypic changes experienced when establishing a colony of *Hermetia illucens* from wild individuals. Phenotypic parameters relevant to mass-rearing that relate to indicators of fitness, fecundity and colony health were measured, including pupae weight, percentage eclosion, post-mating longevity, hatchability of eggs and egg clutch sizes. These traits were monitored over six successive generations, starting from a wild-caught founder population (F0-F5). Hatchability, clutch size and pupae weight are greater in F4 than F1, but a decline is seen from F3 to F4 for all three variables. Post-mating longevity decreased from F1 (100%) to F4 (33%). Most notably, the F5 generation did not develop beyond the pre-pupae stage and the colony collapsed. These phenotypic changes can be explained by a combination of environmental, physiological and genetic effects. The trends observed during earlier generations (F0-F1), may be resultant of the gradual acclimation to artificial environments while natural selection is relaxed. The trends observed at the end of the study period (F3-F4) indicate negative effects of a colony maintenance strategy conducive to inbreeding. The subsequent colony collapse in the sixth generation, despite constant environmental conditions, eludes to possible genetic

effects, which may include inbreeding or genetic drift as a result of relatively small population sizes.

2.1. Introduction

Optimising mass-rearing systems for *H. illucens* is still in its infancy and understanding the phenotypic changes associated with the process of establishing a colony before upscaling production may aid in the efficient management of commercial colonies. Different morphological, physiological and behavioural changes occur from the founding event throughout colony establishment. There is an initial reduction in population size due to the bottleneck at the founder event as a subset of wild individuals is isolated. The population then undergoes multiple selective sweeps through natural selection in captivity, relaxed natural selection and artificial selection (Gilchrist *et al.*, 2012). The pressure to adapt, again reduces the effective population size, causing bottlenecks throughout colony establishment. Another bottleneck may be caused by overestimating measures of fitness, such as hatchability or pupae weight. These parameters are typically used in stocking adult cages and implementing feeding regimes. In the instance of overestimating hatchability, it may be indirectly assumed that the effective population is high (based on egg clutch weights), but with low hatchability, the effective population is in fact decreased. This emphasises the importance of monitoring phenotypic variables throughout colony establishment. Another possible scenario is when natural selection is relaxed by the elimination of for instance, starvation risks, the population may gradually expand, or remain stable. Increasing individual numbers (by implication population density) in a cage, can dramatically influence mating behaviour and development and also act as selective pressures. These effects are variable at different life-stages, affecting larval development time, pupae size and adult fecundity and fertility (Brent, 2010).

Once the population has adapted to the new conditions it may gradually expand and become stable. At this point, the population has been moulded to possess a variety of phenotypic traits conducive to

commercial mass-rearing. These include shortened developmental time, increased fecundity and increased feeding efficiencies (Souza *et al.*, 1988; Kuriwada *et al.*, 2010; Castañé *et al.*, 2014). Documenting the changes of various phenotypes and the interplay between linked phenotypes can aid in the management of populations throughout domestication. Potential bottlenecks can be practically counteracted if they can be identified by phenotypic changes.

The aim of this chapter was to monitor the phenotypic changes which occur when attempting to domesticate a wild population of *H. illucens*. Methodologies describing the mass-rearing of *H. illucens* were described by Sheppard *et al.* (2002) and the establishment and maintenance strategy of the experimental colony was designed to mirror that which is commonly implemented in mass-rearing holometabolous insects (Appendix A: Figure A.1). Phenotypic traits were measured at each generation, namely; clutch size, hatchability, pupae weight, % pupation, % eclosion, sex ratio post-mating longevity (% oviposition). The census population (N_c) was also recorded at each generation to monitor the fluctuations of adult density.

2.2. Methods and Materials

2.2.1. Colony maintenance

Wild pre-pupae (2000 individuals, determined by weight) were collected from Durban, South Africa (29.8587° S, 31.0218° E). Pre-pupae were placed in coarse-grade vermiculite to stimulate pupation. Pupae were placed in a cage constructed from nylon mesh (75 x 80 x 70 cm) (W x L x H). Six discrete, successive generations were maintained in a controlled environment (CE) room. Over the course of the experimental period, environmental conditions were 67.8% (±16%) relative humidity (RH) and 28.2°C (±2°), a photoperiod of 12:12 (Light: Dark) hrs. Adult flies eclosed from pupae inside the cages, dead adults were collected from the cage floor every four days, to allow lekking space for newly eclosed adults and to prevent females from ovipositing amongst dead adults on the cage floor (Tomberlin and Sheppard, 2002).

An outline of the colony maintenance strategy and collection of phenotypic measurements is described in Figure 2.1. Females were provided an oviposition station to attract gravid females. This station consisted of a plastic 1L tub, containing a 300g mixture of 3rd - 4th instar larvae in layer hen feed (Quantum Foods (Pty) Ltd t/a Nova Feeds). The tub was covered with a gauze lid onto which grooved, plastic blocks were placed into which females oviposited. Sheppard *et al.*, (2002) indicated that survival rates are higher when inoculating food sources with neonate larvae, rather than placing eggs directly onto a feeding medium. Therefore, a mass of 0.13g of eggs was collected and placed in an open petri dish. The petri dish was then placed on a gridded plastic mat which covered layer hen feed (60% moisture content). Feed was provided at a ration of 100mg/larvae/day in a plastic tray (40 x 30 x 9cm) (feed recommendations are described by Diener *et al.*, 2009). The gridded plastic mat creates a neutral microclimate in each section of the grid, protecting neonate larvae. Initial feeding was provided for the first six days post-hatching, after which the larvae were fed *ad libitum* in a 44L plastic box (62 x 38 x 18.5cm) until at least 50% of the larvae developed into pre-pupae (Figure 2.1, c). At this stage pre-pupae were gently sieved from remaining dry food and placed in a second 44L plastic box with 1kg course-grade vermiculite. The box was covered with fine mesh and kept sealed with an elastic band to prevent parasitism by the jewel wasp (*Nasonia vitripennis*), while allowing airflow. The box was left undisturbed for two weeks to allow pupation (Figure 2.1, d). Pupae were collected from the vermiculate and placed in a sealed, aerated container until the onset of eclosion. Adult flies were provided clean water every four days through a wick system to prolong adult longevity (Tomberlin *et al.*, 2002) (Figure 2.1, e).

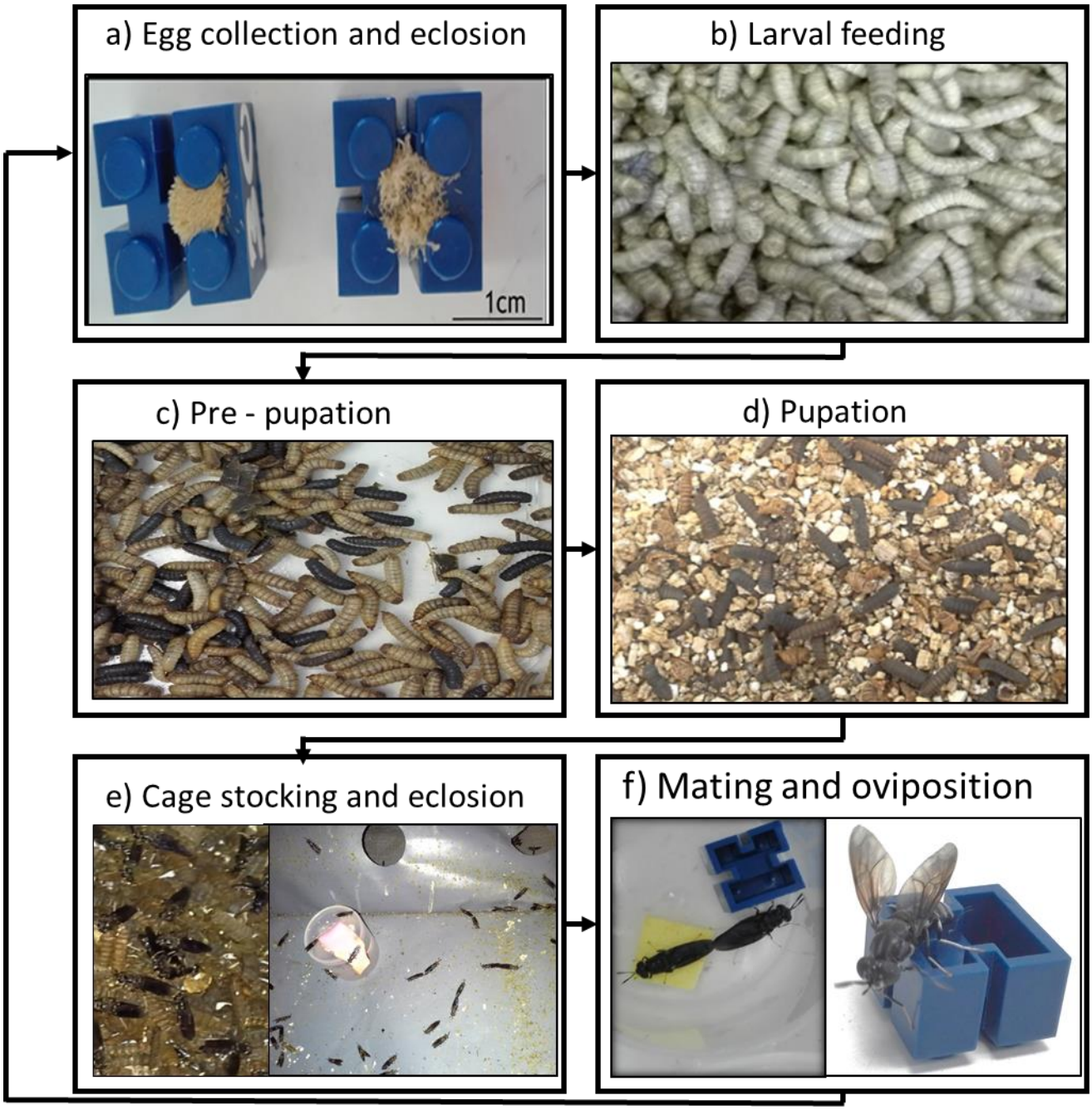

Figure 2.1: An outline of the colony maintenance strategy. Plastic blocks were weighed to determine clutch size. An egg clutch before and during hatching is seen (a). After initial feeding for six days, larvae were moved to a larger container with fresh layer hen mash to continue feeding (b), until pre-pupation (c). Prepupae were placed in vermiculite (d) and covered with fine mesh to allow aeration. Pupae were separated from vermiculite and eclosing adults were released into the cage (e). Adults could mate (e) and were given fresh water every four days with a wick-system. Mating pairs were collected and placed in a 100mL container where the female was left to oviposit in a pre-weighed plastic block (f).

2.2.2. Measurement of phenotypic variables

Clutch size

Mating pairs were placed separately in 100mL clear plastic vials, which were perforated with an equal number of holes to allow equal aeration for each sampled mating pair. A moist sponge ($1cm^2$) and a grooved plastic block was placed in the container to collect eggs. The plastic blocks were weighed before and after oviposition, to determine clutch size by weight, which is the standard unit of measuring clutch size as it correlates to the number of eggs oviposited by each female (Figure B.2).

Hatchability

To understand fluctuations in hatchability irrespective of confounding environmental conditions (Table B.1), temperature and humidity at each generation were recorded at 40-minute intervals (iMonnit™). Hatchability (the percentage of eggs which hatch in each sample) was determined from a random sample of thirty separate clutches. A sample of 100-150 eggs from ten different egg clutches was taken from the oviposition station at three separate occasions for each generation. The sample was then evenly spread to expose all the eggs, which were counted manually with a tally counter, eggs damaged during the spreading process were also counted and deducted from the initial egg count. A moist sponge ($1cm^2$) was placed in the petri dish and the eggs were left to hatch in the same environmental conditions as the adult cage. Unhatched eggs were then deducted from the initial egg count to calculate percentage hatchability.

Pupae weight, % pupation, % eclosion and sex ratio

One week after the onset of pupation a sample of 100 pupae were collected from the vermiculite and weighed individually to four decimal places on an analytical balance (Radwag AS220/C/2). Dead pre-pupae were counted to calculate percentage pupation at the end of each generation. Pupae casings and uneclosed pupae were counted to calculate percentage eclosion (Figure 2.1). To

determine census population size (N_c) and sex ratios, dead adults were stored in 90% ethanol at -20°C and sexed at the end of each generation.

Post-mating longevity (% oviposition)

Throughout the eclosion period, a minimum of 30 mating pairs were collected, until the sample size for clutch weight measurements was n = 30. Of the mating pairs sampled, not all the females survived until oviposition and this was used to estimate female longevity post-mating (represented as % oviposition).

2.2.3. Statistical Analyses

The change of pupae weight, clutch size and hatchability was analysed in STATISTICA.Ink (Statistica v 8.0 Statsoft, Tulsa, OK, USA. 2007). For data which fit a normal distribution as (determined by superimposing a normal curve over a histogram of descriptive statistics generated), Tukey's HSD (honest significant difference) test was performed. As clutch weights did not fit the criteria for Tukey's HSD, a non-parametric Kruskal-Wallis test was implemented to test for significance. The significance of pairwise correlations between pupae weight, clutch weight and hatchability was tested using Pearson's correlation coefficient (r). Significance of correlation of post-mating longevity (described as % oviposition), pupation (%), sex ratio and eclosion (%) in relation to the relevant generation, were assessed using Spearman's (ρ) correlation coefficient.

2.3. Results

Clutch size

Clutch size differed significantly over generations, [H (4, N= 98) = 45.77, ($P < 0.00$)]. A decrease in clutch size was observed in mating pairs collected during F0 and F1, after which an increase was seen from the first captive-bred generation to the next, but then a reduction was seen from F3 to F4, which was not significant (Figure 2.2).

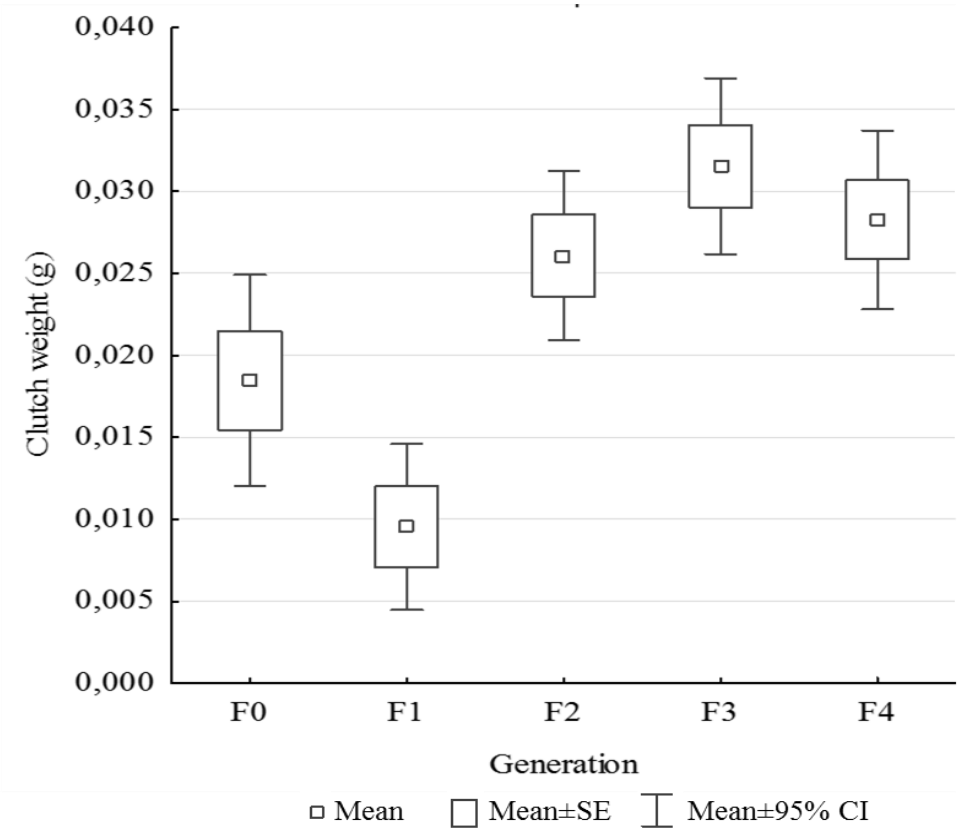

Figure 2.2: The change of average egg clutch weight over generations, starting with F0 (eggs produced by adults collected from the wild) and ending with the fourth captive-bred generation. The size of egg clutches (number of eggs oviposited by each female) was determined by weight. Squares indicate the mean, rectangles indicate the standard error, whiskers indicate 95% confidence interval.

Hatchability

The temperature and relative humidity for eggs collected from the adult of F0-F4 was 26.15±3°, 27.3±0.4°, 28.1±1, 28.4±2° and 30.0±1°C, respectively, and 65.8±12%, 70.8±9%, 69.0±8%, 63.0±8% and 64.3±9% RH, respectively. Hatchability varied significantly over generations ($F = 10.09$, df= 106, $P < 0.00$). Generations F1 and F2 displayed the lowest (mean = 48%; s.d. = 36.81, n = 30) and highest (mean = 91%; s.d.= 7.07, n = 30) percentage hatchability of all generations, respectively. Hatchability decreased significantly from F2-F3 ($P = 0.01$). However, there was no significant difference from F3-F4 ($P = 0.79$) (Figure 2.3).

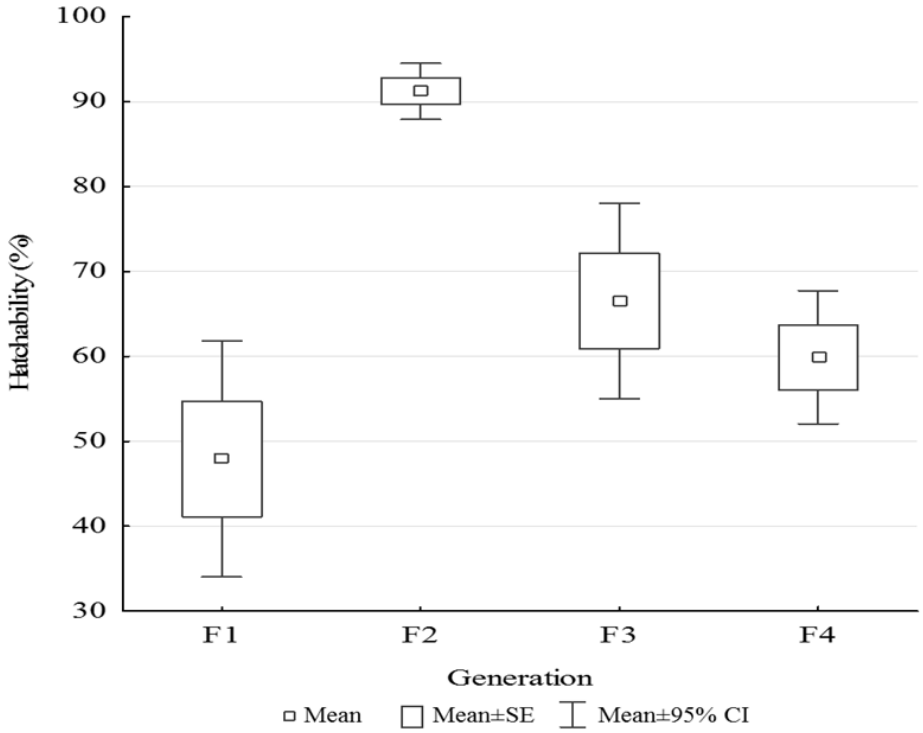

Figure 2.3: Variation in percent of eggs hatching (hatchability) for each generation fluctuated across generation. Thirty samples of eggs were counted before and after hatching. Squares indicate the mean, rectangles indicate the standard error, whiskers indicate 95% confidence intervals.

Pupae weight, % pupation, % eclosion, sex ratio and % oviposition

Mean pupae weights were significantly different for all generations [F (3,400) = 48.40, $P < 0.00$]. The lowest mean weight occurred at F0 (0.13g), while mean weight was the highest at the F3 generation (0.18g) (Figure 2.4)

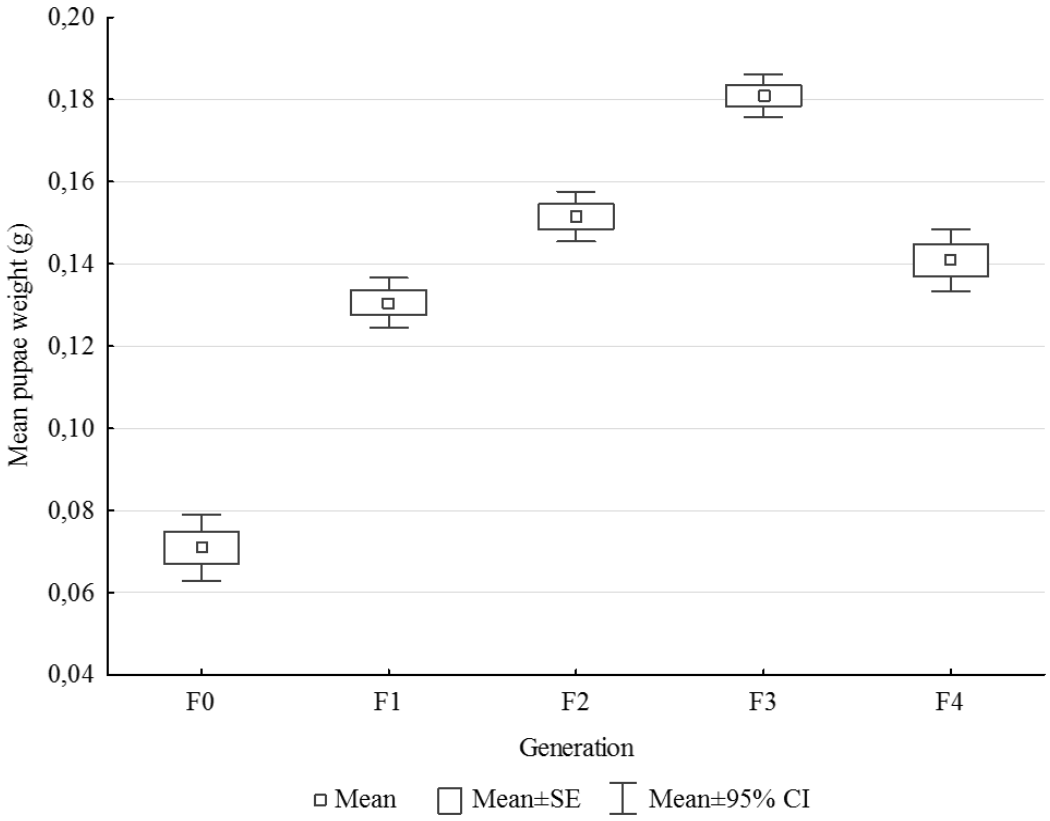

Figure 2.4: Change in mean pupae weight over five generations, namely F0-F4. Squares indicate the mean, rectangles indicate the standard error, whiskers show 95% confidence intervals.

Percentage pupation declined from F0-F1and increased from F1-F4 (F0 = 65.29, F1 = 46.47, F2 = 72.29 F3= 100, F4 = 100). Sex ratios gradually changed in favour of females; F0-36%; F1- 47%; F2- 48%; F3- 46%; F4- 59 %, while the census population (N_c) varied across generations (for generations F0-F4; n = 2434, 484, 3517, 1938 and 689, respectively) (Table 2.1).

Although strong correlations were seen between clutch weight and oviposition (r = -0.80); and clutch weight and pupation (%) (r = 0.94), these relationships were not significant (P = 0.22 and P = 0.07, respectively). Significant correlations were observed between hatchability and clutch weight (r = 0.58; P = 0.02) as well as between oviposition (%) and pupation (%) (r = -0.95; P = 0.05) (Table 2.1).

Table 2.1: The significance of pairwise correlations between commercially-relevant phenotypes was assessed by Pearson's *r*. Correlation coefficients are presented below the diagonal, while significance values (*P*) are presented above the diagonal.

	Pupae weight	Clutch weight	Hatchability	% Oviposition	% Pupation
Pupae weight	-	0.31	0.26	0.42	0.36
Clutch weight	0.58	-	0.02	0.22	0.07
Hatchability	0.29	-0.60	-	0.30	0.21
% Oviposition	-0.47	-0.80	0.47	-	0.05
% Pupation	0.53	0.94	-0.59	-0.95	-

Significant positive correlations were observed in relation to each generation for pupation ($P = 0.05$), while significant negative correlations were observed for post-mating longevity (% oviposition) ($P < 0.05$) (Table 2.2).

Table 2.2: Spearman's Rho (ρ) correlations were drawn to assess the significance of fluctuations in pupation (%), eclosion (%), sex ratio and oviposition (%) in relation to each experimental generation. Significance values are indicated by "*P*" for each generation (F0-F4).

	F0	F1	F2	F3	F4	Spearman's ρ	*P*
% Pupation	65	47	72	100	100	0.87	0.05
% Eclosion	98	63	96	100	98	0.41	0.49
Sex ratio (M:F)	64:36	53:47	52:48	54:46	41:59	0.70	0.19
% Oviposition	89	100	83	53	33	-0.90	0.04

Offspring from F4 remained in the pre-pupae stage and <1% pupated, of which only two individuals eclosed (Figure 2.5).

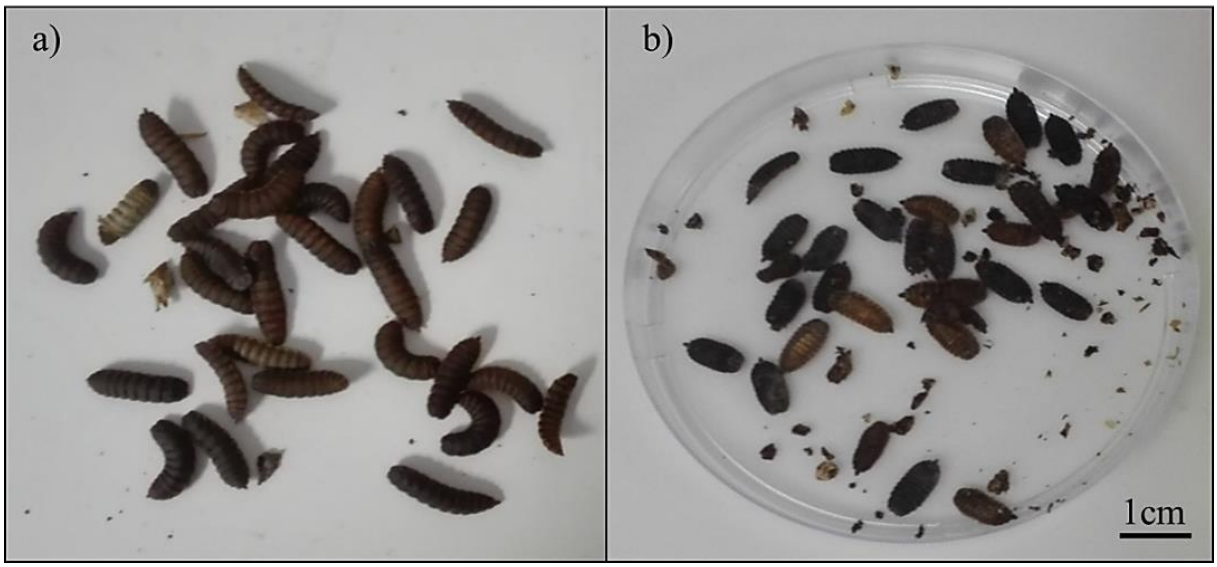

Figure 2.5: The stagnation of F5 in the L6 developmental phase was concerning. Comparison of healthy pre-pupae (a) and the pre-pupae designated to become F5 adults (b). Both samples were photographed at 21days post-hatching.

2.4. Discussion

This chapter aimed at investigating the various phenotypic changes experienced when establishing a colony from wild individuals with the aim of domesticating *Hermetia illucens*. Firstly, quality control measures are noted: the range of temperature and humidity conditions maintained in this study are well within the tolerable ranges reported for *H. illucens* (Sheppard *et al.*, 2002). Furthermore, there is a significant correlation between egg count and egg weight, therefore, clutch weights are interpreted as a measure of total egg production from a single female (Appendix B: Figure B.2).

An initial increase in the size, fecundity and fertility of individuals was seen as the population is introduced to an environment with a homogenous food source and consistent temperature and humidity (F1-F2). During F1-F2 an increase in clutch size was seen. This may have been a manifestation of relaxed natural selection: oogenesis is a nutrient-limited process and a threshold level of resources is necessary for reproduction to begin (Calvo and Molina, 2005). *Hermetia illucens* was fed layer hen feed (as opposed to heterogeneous food in the wild) and sufficient

nutrients were available for oogenesis. Similar results were found during domestication of *Euscepes postfaciatus* and indicate an acclimation to artificial diets in early stages of colony establishment (Shimoji and Miyatake, 2002). Subsequently, clutch sizes decreased, this can be explained by the argument presented by Tomberlin and Sheppard (2002); where adults who fail to mate early enough, will reabsorb oocytes for respiration rather than fertilisation. In addition, the negative correlation between clutch weight and oviposition also provides evidence for this argument. Flies that die before ovipositing have expended all their stored energy, and in the process reabsorbed oocytes. Flies which had an excess of stored energy, lived longer and have larger clutches (*i.e.* they had enough energy to avoid the reabsorption of oocytes). The failure to mate as a result of inadequate lekking space is also discussed in relation to increased pupal size and eclosion. A strong correlation exists between pupae weight, adult weight and clutch size across insect species (Pastor *et al.*, 2011; Smykal *et al.*, 2014). In this chapter, strong correlation was observed between pupae weight and adult size (Appendix B: Figure B.1), while clutch weight and pupae weight showed a weak correlation (Table 2.1). This may be indicative of a shift in resource allocation. According to resource allocation theory, during the domestication process, selection for production-relevant traits causes a shift of resources from reproductive traits (Beilharz *et al.*, 1993). During the rearing scheme applied in this experiment, an inadvertent artificial selection for larger pupae may have occurred by pupae raffling (Gilchrist *et al.*, 2012).

Pupae weights are used when estimating adult densities throughout the cage stocking regime (pupae raffling); they are also an indication of adult size (Appendix B: Figure B.1). A significant increase was seen in pupae weights from F0-F3, while weights slightly declined in F3-F4 (Figure 2.4). Pupa size is determined by larval feeding efficiency; the comparably small pupa size in F0 (0.13g, as opposed to 0.18g in F3) also indicates nutrient assimilation in the wild was lower before the larvae were first introduced to layer hen feed, thereby resulting in relaxed natural selection for individuals

with lower feeding efficiencies. Similar results have been reported in *Bacterocera invadens* during early stages of colony establishment when larvae were provided an optimised feed (Ekesi, 2007).

The effect of increased pupae weights would mean despite having fewer adults in a cage (N_c), larger males would take up more space, resulting in less space to exhibit lekking behaviour. Black soldier fly males exhibit territorial and competitive behaviour toward each other. Aside from the cage walls and floor, cages do not provide lekking sites, which are hypothesised to be vital to the natural competition of males to court females and allow "more fit" males to outcompete weaker males (Tomberlin and Sheppard, 2001). Increased frequency in contact and interaction between larger flies, particularly in this actively lekking species, will increase the metabolic rate of individuals (Pastor *et al.*, 2011). Consequently, the low clutch weight in F1 can also be explained by the low pupae weights in F0 (*i.e.* small pupae will eclose as small adults, which will lay smaller egg clutches). Percentage pupation decreases initially (F0 – F1) and gradually increases from F2- F4, along with percentage eclosion. Similarities in the trends of eclosion and pupation data indicate once the immature stage has committed to pupation, an adult fly will emerge (Table 2.1). This is in accordance with many other studies on BSF and further implies the importance of reaching critical weight (Tomberlin *et al.,* 2002; Holmes *et al.,* 2012). From a commercial perspective, this emphasises the importance of ensuring maximum pupation, while balancing adult densities to facilitate mating behaviour (Singh, 1982; Tchuinkam *et al.*, 2011; Vantomme *et al.*, 2012).

Female fitness is also described by the number of viable offspring they can produce. In this study, egg hatchability is used as a measure of female fertility and fitness. Irrespective of clutch size, hatchability ultimately determines the number of larvae which can be reared to become product. Thus, hatchability is used as an indicator of colony efficiency (egg production), as well as a tool in production management (ensuring adequate amount of food waste is available to sustain larvae). According to data by Holmes *et al.* (2012), hatchability ranges between 65-86% at 70% humidity. In this study, percentage hatchability follows trends of humidity as well (Appendix B: Table B.1,

Figure B.3**).** Similar fluctuations have also been observed in other species during insect domestication (Castañé *et al.,* 2007). The trends in hatchability illustrate the importance of constant environmental conditions in mass-rearing systems, despite acclimation to commercial rearing regimes (Vantomme *et al.,* 2012). Additionally, a significant correlation between hatchability and clutch weight was observed. This correlation was weak (r = 0.58), but is still applicable to mass-rearing *H. illucens*, because it not only indicates a relationship between fecundity (clutch weight) and fertility (hatchability), but can be used to alleviate potential bottlenecks caused by low hatchability rates. In other words, collecting large masses of eggs would increase the probability of hatching and increase the population size.

As previously mentioned, the relationship between pupae weight, adult size and adult longevity has been empirically established in many insect species (Davidowitz *et al.,* 2003). A decline in pupae weight (F3-F4), as well as decreased post-mating longevity are therefore interlinked phenotypes and may also be indicative of a weakening in the ability to assimilate sufficient nutrients to reach critical weight and proceed from the pupal stage. A decrease in the ability to assimilate nutrients at later stages has also been observed in other studies (Fox *et al.,* 2006). Percentage oviposition (post-mating female longevity) showed significant correlation to specific generations using Spearman's ρ. The expenditure in energy as result of increased frequency of contact mentioned previously, could explain the decrease in adult longevity (F3-F4) seen as the flies reach the maximum recorded size (F3) (Pastor *et al.,* 2011). Energy expended in order to lek depletes the accumulated energy stored in the fly's fat body to such an extent that post-mating longevity is decreased. Moreover, a significant negative correlation was observed between pupation and oviposition. This further justifies the notion that an increased frequency of contact (*i.e.* pupation and eclosion are closely linked and more pupae would eclose as flies and increase adult densities) causes an expenditure of stored energy, which decreases the fly's longevity.

The most notable result observed in this study is the stagnation of F5 in the pre-pupae state. Similar observations were made by Holmes *et al.* (2012), where hollow body cavities were observed in dissected pupae, indicating exhaustion of fat bodies. This also delayed time to pupation, which is further anticipated based on results by Smykal *et al.* (2014) and Noriega, (2014). This is in accordance with the notion that critical weight is reached during L5 and that pupariation, is final stage of development at which the accumulation of nutrients suffices to advance to the next developmental stage (Davidowitz *et al.*, 2003). Possibly as a result of inbreeding or genetic drift, the ability to efficiently assimilate nutrients from the layer hen feed was lost, or the appropriate endocrine signalling was hindered which would result in the continued metamorphosis from pre-pupae to pupae. The increased risk of starvation in inbred populations of *Drosophila melanogaster* has also been demonstrated by Kirstensen *et al*. (2005). Inbreeding may have occurred by the reduction in effective population size which can be represented by the reduction in clutch weight and hatchability in the final generations.

2.5. Conclusion

The decrease in phenotypic variables measured in early generations, and again at the final generations in this study show similar trends as those described in multiple investigations of insect domestication. The results presented here are likely an exaggeration of the processes undergone in larger populations, but can be applied to predict possible outcomes in captive colonies. One possible application of these results to commercial rearing is to providing adequate lekking space for males, alleviating the exhaustion of energy reserves by the above-mentioned behaviour. Moreover, the investigation of solely phenotypic effects of establishing a captive colony from wild individuals is limiting. To further understand the observed phenotypic changes, it is recommended that investigation of genetic changes be coupled with this research. Moreover, the initial genetic diversity and relatedness of the founding population is not certain, neither are the maternal effects

on various traits measured in this study. This emphasises the need to develop the molecular tools, specific to BSF to further elucidate the processes of mass-rearing.

Chapter 3: Development and characterisation of microsatellite markers for the black soldier fly, *Hermetia illucens* (Diptera: Stratiomyidae)

Abstract

With the increased application of molecular markers in population genetics studies, a wide variety of techniques have become available for relatively easy and cost-effective isolation and development of molecular markers. Microsatellite markers possess a variety of characteristics which make them ideal to study the changes of genetic diversity associated with commercial colony establishment. To aid in the understanding of phenomena occurring during domestication and mass-rearing *Hermetia illucens*, this chapter reports on the development and characterisation of novel microsatellite markers for *Hermetia illucens*. From the draft genome assembly, a set of 40 tetra-nucleotide microsatellite loci were identified and primers for marker amplification were designed. Of the 40 primer pairs, 29 were successfully optimised and were fluorescently labelled for polymorphism screening. The annealing temperatures of the labelled primers were readjusted and 10 of the 29 markers consistently amplified, while also showing polymorphism. To furthermore assess the utility of the markers, three multiplex-PCR reactions were developed and 37 wild individuals genotyped. The number of alleles for each locus ranged 4 - 21. Polymorphism information content ranged from 0.52 to 0.90, while observed and expected heterozygosity ranged from 0.30-0.65 and 0.55-0.91, respectively. The newly developed microsatellite markers proved useful in individual identification and parentage analysis. The markers reported in this chapter will be valuable for assessing the levels of genetic diversity for commercial *H. illucens* colonies.

3.1. Introduction

Microsatellite loci are simple sequence repeat loci, consisting of two to six nucleotides repeated in tandem, in tracts of <1kb. Their codominant nature allows for the discrimination between heterozygous and homozygous genotypes. They are also ubiquitous throughout the entire genome of almost all living organisms, making them valuable tools to study a wide range of taxa with application to a wide range of studies. The abundance of microsatellites occurring within the genome, their polymorphic nature and ability to be analysed using PCR make them a favourable option for initial investigations into the genetic consequences of BSF mass rearing. In addition, incorporating multiple markers in multiplex reactions, allows for high throughput, cost-effective assessment of genetic diversity (Chambers and MacAvoy *et al.*, 2000; Chistiakov *et al.*, 2005; Webster and Reichart, 2005).

Microsatellite markers have previously been utilised to study the process of insect domestication and marker assisted selection to improve production efficiencies (Nagaraju, 2002; Kuriwada *et al.*, 2010). However, prior to application, a careful process of marker design, development and characterisation must be undergone. Extensive baseline research is required to confirm the utility of microsatellite markers, which can be assessed by the number of alleles (N_A), observed heterozygosity (H_O), expected heterozygosity (H_E) and polymorphism information content (PIC).

When establishing new colonies from the wild or pre-existing colonies, it is important to document the genetic diversity of populations. This aids not only in ensuring that the widest range of alleles are brought in as part of the founding population, but also allows for the implementation of management strategies to maintain a stable gene-pool within the facility over multiple generations. Currently, the development of molecular tools specific to *H. illucens* genetics is in its infancy. To facilitate the expansion of knowledge regarding BSF genetics, a panel of microsatellite markers was developed and characterised. This was approached by screening a *de novo* genome assembly for

tetra-nucleotide repeats (these repeat units decrease the occurrence of stutter profiles) (Miller and Yuan 1997; Mueller and Wolfenbarger, 1999). Primers were designed for selected repeat sequences and optimised for application in a multiplex panel for subsequent genotyping and genetic diversity estimates.

3.2. Methods and Materials

3.2.1. Sample collection and Genomic DNA preparation

A sample of 37 adults (18 males and 19 females) originating from Durban, South Africa, was stored in 90% ethanol at -20°C. Head and thorax tissue was removed with a sterile surgical scalpel and dried on a heating block at 40°C immediately before genomic DNA extraction. Genomic DNA was prepared according to the cetyltrimethyl ammonium bromide (CTAB) protocol described by Saghai-Maroof *et al.* (1984), with concentrations of extraction buffer reagents modified to accommodate extraction from insect tissue (Wang and Wang, 2012). Nucleic acid quantification was carried out on a NanoDrop ND-1000 spectrophotometer v.3.0.1 (*NanoDrop*). Prior to (PCR) analysis, each DNA sample was adjusted to a working concentration of 50-100ng/µl and stored at -20°C. Genomic DNA from these samples was used throughout the marker optimisation and characterisation process.

3.2.2. Primer design

Paired-end (PE) reads from a whole genome sequence of male (accession: SRX265066) and female (accession: SRX265065) *Hermetia illucens* were downloaded from the sequence read archive (SRA) in GenBank. The SRA files were converted to FastQ format using the SRA Toolkit (v2.3.4-3). The male had a total of 32,383,756 PE reads, and the female 30,424,840 PE reads. *De novo* assembly was performed using CLC Genomic Workbench (v7.0.4) on each of the sexes, separately (Table C.1). The option to create contigs without having to map the reads back was selected. A k-mer size of 24 nucleotides for the combined assembly and 23 nucleotides for both the male and

female assembly was automatically calculated by the assembler. Bubble size was set to 100 bp and all contigs shorter than 500 bp were removed in each of the assemblies. In combination, male and female reads were used to produce a single genome assembly. A total of 498 polymorphic tetra-nucleotide markers were identified after stringent criteria were implemented for returning microsatellites in MSATcommander (Faircloth, 2008). These criteria included screening for tetra-nucleotides with a minimum of four repeats with at least two C's or G's, to avoid ATTT-like loci. From this data, 40 tetra-nucleotide microsatellites had sufficient flanking sequences for primer design**.** Specific primers were designed on Primer3 (Untergasser *et al.*, 2007). Primers had to be at least 50bp from the microsatellite locus and possess expected product sizes and annealing temperatures which would allow for downstream multiplex compilations.

3.2.3. PCR optimisation

All PCR reagents were sourced form KAPA Biosystems, Cape Town, SA (KAPA Taq PCR Kit; KR0352). The optimisation strategy from initial PCR optimisation to the compilation of multiplex panels is depicted in Figure 3.1. A touchdown PCR (TD-PCR) approach was initially implemented, whereby the annealing temperature (T_a) was gradually decreased during the cycling process. The TD-PCR approach was carried out prior to gradient PCR, in order to target a 5°C T_a range in a single reaction (Hecker and Roux, 1996). At the beginning of the cycling stage, the T_a was set either 5°C above or below the recommended melting temperature. Melting temperatures were tested as recommended by primer manufactures (Whitehead Scientific, Cape Town, SA). TD-PCR cycling conditions were run as per Laborda *et al.* (2009); thermocycler parameters were 94 °C for 2 min; 2x (10 cycles of 94°C for 1 min, recommended T_a (-1°C/cycle) for 1 min and 72°C for 2 min); then denaturation at 18 cycles of 94°C for 1 min, annealing at T_m for 1 min and extension at 72°C for 2 min; and final extension of 72°C for 5 min. Reactions were performed in 10-µl reactions with ~50-100ng DNA, 1X Buffer (containing 1.5mM $MgCl_2$), 0.2mM of each dNTP, 0.2µM of each primer and 1 U Taq polymerase. Initially, fragments were not successfully amplified and $MgCl_2$

concentrations were readjusted by increasing $MgCl_2$ concentrations in increments of 0.25mM to make up the final $MgCl_2$ concentration.

After TD-PCR, four to six individuals were used for initial singleplex PCR tests, to specify T_a. Briefly, thermocycler parameters were: initial denaturation at 95°C for 3min (1 Cycle); denaturation at 95°C for 30sec, annealing at relevant T_a for 30sec, extension at 72°C for 30sec (35 Cycles); final extension at 72°C (1 Cycle) and hold at 4°C. Reactions were performed in 10-µl reactions with ~50-100ng DNA, 0.2mM of each dNTP, 0.2µM of each primer, 1 U Taq polymerase, 1X Buffer A (containing 1.5mM $MgCl_2$); $MgCl_2$ was supplemented to a final concentration of 2.0mM.

Markers with consistent amplification were identified by electrophoresis on ethidium bromide-stained agarose gel (1% w/v EtBr; 2% agarose; 1 X TBE). Negative controls wherein template DNA was omitted were added to control for contamination. Promega 100 bp molecular size ladder was used for preliminary size determination. Multiplexes were preliminarily considered based on annealing temperature and expected allele range, this served as the criteria for selecting PET, FAM, NED and VIC dye-labels for the forward strand primers (Whitehead Scientific, Cape Town, SA). Labelled markers were then screened for polymorphism by combining single reactions in poolplexes prior to capillary electrophoresis on ABI3730xl Genetic Analyser™ (Life Technologies) with GeneScan™ 500 LIZ® (Life Technologies, Foster City, CA, USA) as an internal sizing standard. Polymorphic markers were identified in GeneMapper® v4.1 (Life Technologies). Finally, three new multiplex PCRs were considered based on annealing temperatures, expected allele range and fluorescent labels of the markers which were polymorphic (Figure 3.2; Table 3.3).

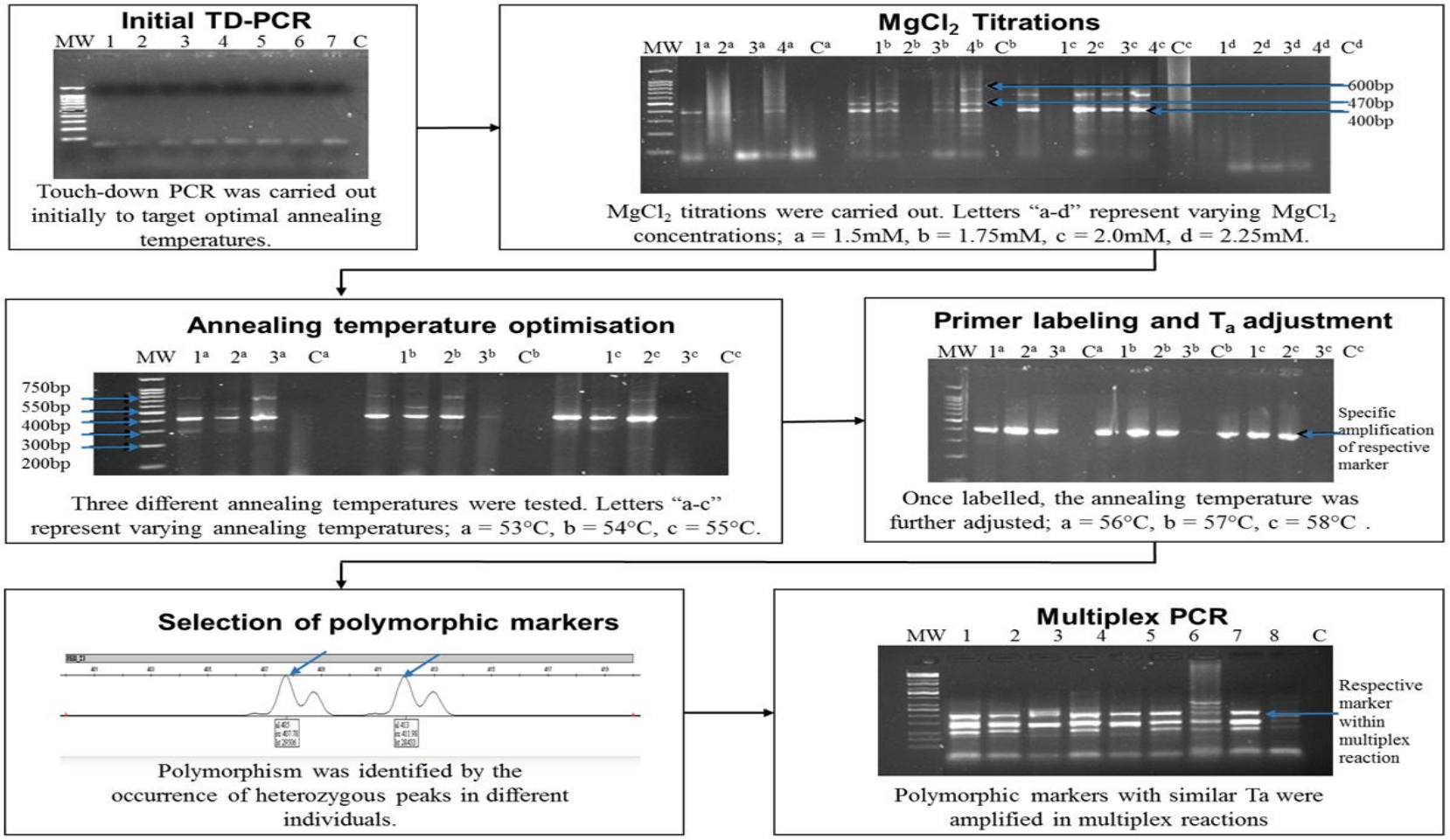

Figure 3.1: Number indicate sample individuals tested, while "C" indicates a negative control, in which template DNA was omitted from the PCR reaction. TD-PCR was tested at -5'C from the recommended T_a, then at 53-55'C. The marker was then labelled with NED and T_a readjusted up to 58'C. The marker was identified as polymorphic and incorporated into a multiplex reaction (Panel 2).

Multiplex amplifications were performed in 10-µl total reaction volumes containing KAPA2G Fast Multiplex PCR Kit (KAPA Biosystems, Cape Town, SA); 0.8 µM of each primer and ~50-100ng DNA. Reaction parameters were as follows: initial denaturation at 95°C for 5min, 35 cycles of 95°C for 15 sec, T_a for 30 sec, 72°C for 50 sec, and a final extension at 72°C for 90 sec. Negative controls were again included in all amplifications prior to capillary electrophoresis on ABI3730xl Genetic Analyser™ (Life Technologies) with GeneScan™ 500 LIZ® (Life Technologies, Foster City, CA, USA) as internal standard.

3.2.4. Marker validation and characterisation

Alleles were scored in the software GeneMapper® v4.1 (Life Technologies). MICRO-CHECKER v.2.2.3 (Van Oosterhout *et al.*, 2004) was used to detect genotyping errors due to the presence of null alleles, allele stuttering and large allele drop out. Null allele frequencies were estimated using the Brookfield 2 method (Brookfield, 1996). Genotypic frequencies were tested for conformation to Hardy-Weinberg Equilibrium (HWE) by Fisher's exact probability test in the web version of GENEPOP (10 000 dememorisation, 500 batches, and 5000 iterations per batch) (Raymond and Rousset, 1995; Rousset, 2008). Locus-specific F_{IS} was estimated by the Weir and Cockerham (1984) method in GENEPOP. The number of alleles (N_A), number of effective alleles (A_E), observed heterozygosity (H_O), expected heterozygosity (H_E) and polymorphism information content (PIC) were calculated in Microsatellite Toolkit v3.1.1 (Park, 2001). Allelic richness (R_S) was calculated by rarefaction in HP-RARE1.0 (Kalinowski, 2005). Allele frequencies and additional measures of marker utility, in individual identification and parentage analysis were calculated in GenAlex v.6.41 (Peakall and Smouse, 2006): the probability of identity (P_{ID}), probability of sibling identity (P_{IDsib}) as well as the probability of exclusion when one or two parents are known (P_1 and P_2). The significance of linkage disequilibrium between marker combinations was assessed using Fisher's exact test in GENEPOP.

3.3. Results and Discussion

Statistic for the draft genome assembly are listed in Table C.1 (Appendix C). Of the 40 primer pairs, 29 were successfully optimised in PCR reactions and ten were finally identified as polymorphic and showed consistent amplification. Markers were used in three multiplex reactions (Figure 3.2), based on allele size range and annealing temperatures, determined during marker optimisation (Figure 3.1). Marker characteristics and validation estimates are summarised in Table 3.1.

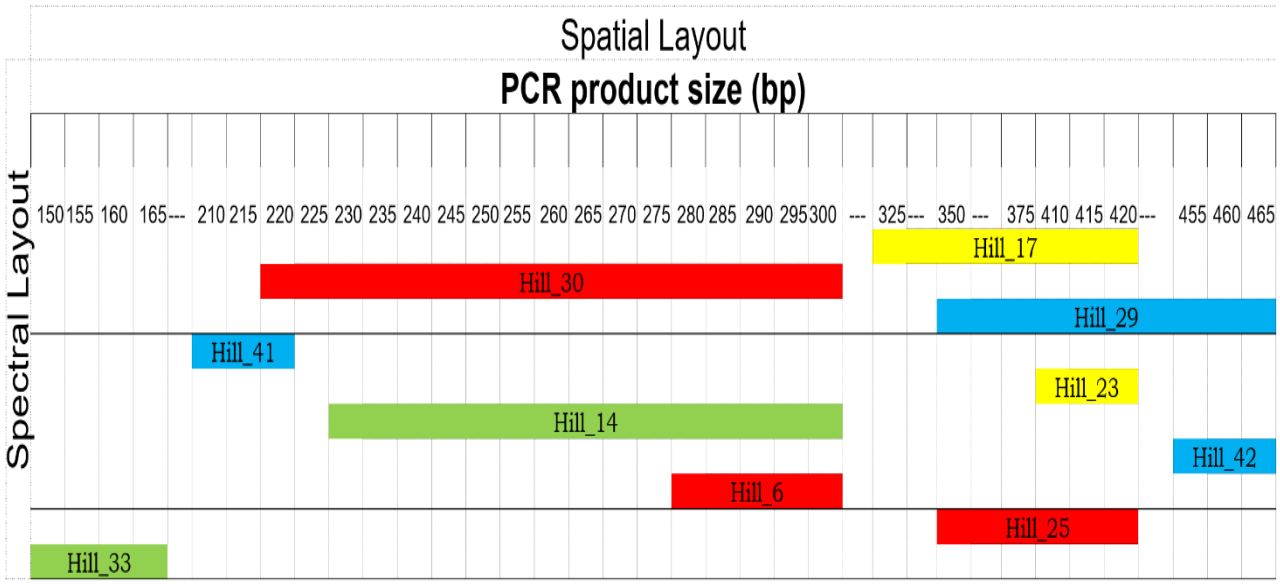

Figure 3.2: The spatial and spectral layout of three multiplex panels for *Hermetia illucens*. Markers were arranged according to T_a, expected size and fluorescent labels. Fluorescent dyes are represented- blue: FAM; red: PET; yellow: NED; green: VIC. Table 3.3 describes the characteristics of these markers.

Table 3.1: Characteristics and genetic diversity summary statistics of ten microsatellite markers for *Hermetia illucens*, arranged in three multiplex panels. Fluorescent labels NED, PET, VIC and FAM are indicated on the forward primer; N = sample size, N_A = number of alleles, A_E = number of effective alleles; R_S = allelic richness; H_O = observed heterozygosity, H_E = expected heterozygosity, F_{IS} = fixation index, PIC = polymorphism information content, HWE = significance of deviation form Hardy-Weinberg equilibrium, F_{NULL} = null allele frequency (Brookfield, 1996), P_{ID} = probability of identity, P_{IDsib} = probability of sibling identity, P_1 = exclusion probability knowing one parent; P_2 = exclusion probability knowing both parents.

Locus name	Repeat motif	T_a (°C)	Size range (bp)	Primer Sequence (5'-3')	N	N_A	A_E	R_S	H_O	H_E	F_{IS}	PIC	HWE	F_{NULL}	P_{ID}	P_{IDsib}	P_1	P_2
Panel 1																		
Hill_17	(CCGA)n	60	316-432	F: NED-GTCAGTATCCCAGGGCACC R: TCGTTGCATTGAAGTCGCC	37	10	3	3	0.32	0.7	0.5445	0.67	0.0000	0.2679	0.12	0.43	0.50	0.31
Hill_30	(ACAG)n	60	208-428	F: PET-CCTAGGCTCCGAAATAGCC R: GCGATCGAGTGCTGCAAG	34	21	11	4	0.65	0.91	0.3026	0.9	0.0000	0.2942	0.01	0.30	0.82	0.70
Hill_29	(ACTT)n	61	303-583	F-AGGGCGATCACAGTTCCAG R-TTCCCGCGCTTACCGAATG	36	21	7	3	0.67	0.85	0.2558	0.84	0.0000	0.1636	0.03	0.33	0.73	0.58
Panel 2																		
Hill_41	(ATCC)n	58	205-221	F: FAM-AATCACGACGGGTAAACGG R: AGCATTTGCCGAGAGCATC	37	4	2	2	0.3	0.59	0.5038	0.52	0.0000	0.2395	0.24	0.52	0.32	0.18
Hill_23	(ACGG)n	58	405-417	F: NED-CCTAGGCCGAATTTGCTTCC R: AACGAAATGCAGGTAGGCG	35	4	3	2	0.57	0.62	0.0951	0.57	0.3459	0.2104	0.20	0.49	0.37	0.21
Hill_14	(ACAG)n	58	210-298	F: VIC-TCCGGCTTGTTTGAATACCG R: ACTTTGCTAGTAATTGTGGGC	35	8	5	3	0.46	0.8	0.4400	0.77	0.0000	0.3142	0.07	0.37	0.61	0.43
Hill_42	(ACCT)	58	450-502	F: FAM-AGGTTACGAAGAAGGGCCG R: GAGCCGAATTTGTTGGGCG	32	9	5	3	0.5	0.81	0.3754	0.78	0.0000	0.3565	0.06	0.36	0.63	0.45
Hill_6	(AGAT)n	59	230-302	F: PET-TGGCCTCTTCTTCCAAGTG R:ACAGGTTTCAGGAAAGGTTAGAAG	37	10	4	3	0.41	0.75	0.4899	0.73	0.0000	0.2069	0.08	0.39	0.57	0.38
Panel 3																		
Hill_25	(CCGT)	54	359-439	F: PET-CCTCCCATCCAAGGTCTCG R: GTGCGATCAATGGCCTACG	35	13	5	3	0.51	0.79	0.3834	0.78	0.0000	0.5533	0.06	0.37	0.63	0.45
Hill_33	(GGAT)n	54	153-173	F: VIC-GGCATCAATTCCCTCTATGCG R: GTGCCAGCTGTAACCGAAC	36	6	2	2	0.39	0.54	0.0951	0.52	0.0022	0.5353	0.24	0.54	0.34	0.17

The informativeness of a marker is a function of marker polymorphism, *i.e.* the number of alleles detected as well as their frequency distributions. Average PIC for all markers was 0.71 and ranged 0.52 – 0.90. The number of effective alleles (A_E) was less than half the number of alleles (N_A) for all loci (2 - 11 and 4 - 21, respectively). This discrepancy between A_E and N_A, indicates that there may be a number of lower frequency alleles. In addition, loci with lower numbers of alleles such as Hill_41 (N_A = 4), tend to have lower heterozygosity values than loci with a higher number of alleles such as Hill_30 (N_A = 21), (H_O = 0.30 and H_O = 0.65, respectively). Overall, the observed and expected heterozygosity ranged 0.30 – 0.67 and 0.54 – 0.91, respectively. This broad range in H_O and H_E results from the large variation of N_A per locus, and large allele frequency distributions. Allele frequency distributions are shown in Figure C.1 (Appendix C).

Significant deviations from Hardy-Weinberg equilibrium were evident in eight loci ($P < 0.00$). This was attributed to homozygote excess for all these loci, which was also evident in positive F_{IS} values (Table 3.1). Overall, analysis did not indicate evidence of genotyping errors being attributed to large allele dropout or stuttering. However, a possible cause of scoring error may be attributed to null alleles, which were detected in high frequency for most of the markers. Null allele frequencies ranged between 0.16 and 0.53 using the Brookfield 2 method (Brookfield, 1996) (Table 3.1). The occurrence of null alleles leads to a detection of homozygotes during genotyping, and decreases observed heterozygosity. Insects generally show high instances of null alleles, along with other taxa which are highly fecund, such as molluscs (Chapuis and Estoup, 2006). Departure from HWE may also be a result of the animal's reproductive biology, rather than null alleles in the dataset. *Hermetia illucens* is a docile fly which lays egg clutches of 320-620 eggs (Tomberlin *et al.*, 2002). This high fecundity, coupled with the low dispersal range of BSF adults would result in an increased probability of sampling related individuals - increasing calculated homozygosity.

Nevertheless, consequences of null alleles regarding downstream applications were not overlooked. Null alleles could lead to an overestimation of allele frequencies, inflation of relatedness estimates,

as well as skewing probability of identity and exclusion probabilities in parentage analyses (Dakin and Avise, 2004). As such, the utility of markers in individual identification and parentage analysis were calculated. The P_{ID} and P_{IDsib} ranged 0.01-0.24 and 0.30-0.54, respectively. This is less significant than what was found in other studies using similar sampling strategies, for insects with similar life-histories (Rabeling *et al.,* 2013; Aarens *et al.,* 2015). However, when the markers were assessed in combination, the power of markers to be used in individual identity and sibling identity increased greatly ($P_{ID} = 1.2X10^{-11}$ and $P_{IDsib} = 1.1 \times 10^{-4}$). In addition to this, exclusion probabilities based on knowing the genotype of either one or both parents using markers independently, was low (P_1 = 0.32-0.82 and P_2 = 0.17-0.70, respectively), while the marker utility increased greatly when the panel was used in combination ($P_1 = 1$ and $P_2 = 1$). This emphasises the importance of using microsatellite markers in combination when determining probability of identity and parentage analysis (Woods *et al.,* 1999; Waits *et al.,* 2001). Probability of identity and exclusion probabilities relative to the combination of markers are presented in Figures C.2-4 (Appendix C). Moreover, the markers presented here can be used to investigate adult mating behaviour. A monogamous mating strategy is assumed in BSF, based on field and laboratory observations (Copello, 1926; Tomberlin and Sheppard, 2001). However, BSF is a lekking species, and the possibility of multiple copulations cannot be ignored. Mating frequency has implications for the effective population size, genetic structure of the population and maintenance of genetic variation (Zouros and Loukas 1989; Zeh *et al.,* 1997). *Bactrocera cacuminata,* for example, is a species of fruit fly which exhibits lekking behaviour and multiple paternity. Understanding mating rates in this species has aided in the management of adult colonies to conserve genetic diversity in commercial colonies (Song *et al.,* 2007).

To validate the application of these markers, the significance of linkage (or gametic) disequilibrium between marker pairs was assessed. Markers in significant linkage disequilibrium would result in type I errors due to pseudo-replication of genotypic data (Selkoe and Toonen, 2006). The statistical

significance of linkage disequilibrium (LD) was tested for 45 pairs of marker combinations by Fisher's exact test. Out of these combinations, two pairs (4%) were in significant LD ($P < 0.01$). Marker combinations which had significant LD at the 1% nominal level were Hill_42 with Hill_6 and Hill_17 with Hill_33. The small percentage of LD may be a random result of sampling and the utility of these markers in assessment of genetic diversity is further confirmed.

3.4. Conclusion

The markers described here will be useful in studies concerning the evaluation of genetic diversity, individual identity and parentage analyses - all of which can be implemented in designing colony management strategies. These markers present high variability and are therefore useful in estimating genetic diversity throughout colony establishment. Furthermore, the number of alleles per locus reported in this study is likely less than the true number of alleles in the total wild population. Estimates of the number of observed alleles in a population can be affected by the number of individuals sampled, thus influencing the true allele frequency from one population to another. By sampling a larger number of individuals, it is likely that a larger number of alleles and effective alleles would be detected, increasing PIC and power of the markers in downstream applications. Moreover, PIC values are comparable for different genetic markers and allow for the implementation of different of markers to address a broad range of research questions relevant to colony establishment. An example of this is assessing the diversity of BSF populations globally. This would assist in selecting a genetically diverse founding population for initial colony establishment. In addition, the occurrence of low frequency alleles is important to note, as they can be lost during random genetic drift. Gaining insights such as this, illustrate the value of using microsatellite markers in implementing colony management strategies which conserve genetic diversity.

Chapter 4: Assessment of genetic diversity in black soldier fly, *Hermetia illucens* (Diptera: Stratiomyidae), during colony establishment

Abstract

To successfully establish a commercial colony, the effect of isolating individuals from the wild and rearing them in an artificial environment must be investigated. A wide range of demographic forces are at play which increase the probability of inbreeding and genetic drift and cause declines in genetic diversity. The decline in genetic diversity from a wild (F0) to an F5 generation of the black soldier fly (BSF), *Hermetia illucens*, was assessed using microsatellite markers. Three multiplex PCR panels with a total of ten microsatellite markers were used to assess the genetic diversity across six successive generational cohorts. Genetic diversity analyses showed a decline in diversity from the wild founding population to the final generation, with both allelic richness and observed heterozygosity declining from 2.84 to 2.34, and 0.48 to 0.31, respectively. Indicators of population bottlenecks were evident in the estimation of effective population size. Significant differences were noted between all cohorts using pairwise-F_{ST} analysis and a large differentiation is seen from F0 to F5 (0.186, $P < 0.01$), with the greatest difference observed between F3 and F5 (0.343). Phenotypic parameters which significantly correlated to changes in genetic diversity were pupation (%), clutch size, pupae weight and female post-mating longevity ($P < 0.05$). Results indicate small N_e resulted in the amplification of inbreeding depression and genetic drift. All genetic diversity estimates were compared to a sample from a currently maintained commercial population (F28), which was significantly lower in observed heterozygosity, and significantly differentiated from experimental generations. This study illustrates the bottlenecks causing a loss of genetic diversity that can occur during domestication, when genetic management of the populations is not implemented.

4.1. Introduction

The black soldier fly, *Hermetia illucens*, is currently reared at a mass-scale in a small number of facilities, globally. As the industry rapidly develops and gains momentum, it is expected that there will be a substantial expansion in the number and size of facilities mass-rearing *H. illucens*. Currently, the main technique of starting colonies entails either collecting wild individuals or outsourcing a founding population from smaller facilities. This may have severe implications for the successful establishment of new colonies and affect colony health at later stages. As the industry develops, eventually strategies for production performance will be implemented, such as selective breeding. Before considering selective breeding practices, the successful establishment of a colony must first be mastered. As an industry in its infancy, BSF mass-rearing has the advantage of learning from the mistakes of traditional animal breeding regimes. The effects of domesticating wild insects and their adaptation to artificial mass-rearing environments are widely studied in species of economic importance (Shimoji and Miyatake 2002; Gilchrist *et al.,* 2012; Baeshen *et al.,* 2014).

The first bottleneck which is overlooked early in colony establishment occurs during the founding of the captive population. The founding of a comparatively small population from the wild restricts the number of individuals which would effectively contribute to the subsequent generation (Baeshen *et al.,* 2014). This results in the founder effect and entails a loss alleles due to the small effective population size of the newly isolated population. Highly fecund species, such as *H. illucens*, are particularly at risk of the deleterious consequences of the founder effect. A founding population of 2000 individuals may represent a contribution of only eight mating pairs (Tomberlin, 2002). This increases the chances of losing rare alleles by random genetic drift and results in an overall decrease in heterozygosity. Moreover, the probability of inbreeding is increased by isolating a small population of a species with low dispersal capabilities. *Hermetia illucens* is a docile fly and

its entire life-cycle occurs in close proximity to food sources (Tomberlin and Sheppard. 2002). Therefore, establishing a colony from one location will automatically limit genetic diversity.

Following the founding event, the commercial population will undergo multiple selective sweeps, the first of which would result in a decrease in population size through natural selection in captivity. Simultaneously, certain aspects of the artificial rearing environment such as the provision of optimal feeds *ad libitum* will result in relaxed natural selection. In optimised mass-rearing programmes, artificial selection is intentional. However, unintentional artificial selection may take place as well and has been documented in typical strategies for mass-rearing holometabolic insects (Gilchrist *et al.*, 2012).

The interplay of the aforementioned interacting factors can lead to either colony collapse or specific adaptation and domestication (Price and King, 1968; Bartlett, 1993; Baek *et al.*, 2014). The aim of this study was thus to assess the change of genetic diversity throughout colony establishment. A colony was founded from a collection of wild individuals (n = 2000) and reared for six successive, discrete generations. A colony management scheme common in holometabolous insect rearing was implemented (Chapter 2). Individuals were sampled at each generation constituting six cohorts, while an additional sample from an existing commercial population was also included. Individuals were genotyped using ten microsatellite loci. Genetic diversity was then assessed across all generations. The relationships between genetic diversity and commercially-relevant phenotypes were also assessed.

4.2. Materials and Methods

4.2.1. Colony maintenance and sampling

The experimental colony was founded from 2000 wild individuals originating from Durban, South Africa (29.8587° S, 31.0218° E). Colony maintenance and rearing protocols are described in

Chapter 2. To establish every successive generation, 0.13g of eggs (± 40 000 eggs) was collected from the preceding generation. Eggs were sourced from as many different clutches as possible.

For genetic analyses, six experimental cohorts, representing the six consecutive generations after the establishment of a founder population (F0-F5), and one commercial cohort (F28) were sampled. At the time of sampling, the commercial population was in the 28th generation after establishment. This colony was established in 2012 from a wild founding population (Durban, South Africa). Since the founding of this population, no supplementation has been implemented and the size of the founding population was also not recorded. For each experimental generation, 15 mating pairs were collected and served as the source of genomic DNA material for this study, (n = 30 in total; 15 of each sex). The final generation, F5, stagnated in the immature larval stage (*i.e.* adults did not eclose), therefore, 30 pre-pupae were sampled to represent this generational cohort. A random sample of 30 adults was also collected from a commercial population of BSF (F28). All samples were stored individually in 90% ethanol at -20°C.

4.2.2. DNA preparation and Microsatellite genotyping

Genomic DNA was prepared as described in Chapter 3. Thirty individuals (n = 15 of each sex), for each generational cohort (210 individuals in total), were genotyped using ten polymorphic tetra-nucleotide microsatellite markers (described in Chapter 3). Multiplex PCR amplifications were performed in 10µl total reaction volume containing KAPA2G Fast Multiplex PCR Kit (KAPA Biosystems, Cape Town, SA), 0.8 µM of each fluorescently-labelled primer and 50-100ng of template DNA. Reaction parameters were set as follows: initial denaturation at 95°C for 5 min. 35 cycles of 95°C for 15 sec., Ta for 30 sec. 72°C for 90 sec. and a final extension at 72°C for 30 sec. (1 cycle). Aliquots of PCR reactions (4µl) were visualized by agarose gel electrophoresis to determine successful amplification. Capillary electrophoresis was conducted at the Central Analytical Facilities in Stellenbosch on the ABI3730xl Genetic Analyser™ (Life Technologies); GeneScan™ 500 LIZ® (Life Technologies, Foster City, CA, USA) served as an internal sizing

standard. Alleles were scored in the genotyping software package GeneMapper® v4.1 (Life Technologies). MICRO-CHECKER v.2.2.3 (Van Oosterhout *et al.*, 2004) was used to detect genotyping errors due to the presence of null alleles, allelic drop out and stuttering. Null allele frequencies were estimated using the Brookfield 2 method (Brookfield, 1996).

4.2.3. Genetic diversity statistics

Genetic diversity statistics were computed for each cohort. Genotypic frequencies were tested for conformation to Hardy-Weinberg Equilibrium (HWE) by Fishers exact probability test in the web version of GENEPOP (10 000 dememorisation, 500 batches, and 5000 iterations per batch) (Raymond and Rousset, 1995; Rousset, 2008). The number of alleles (N_A), observed heterozygosity (H_O), expected heterozygosity (H_E) and unbiased expected heterozygosity (uH_E) were calculated in the Excel Microsatellite Toolkit v3.1.1 (Park, 2001). Additional genetic diversity statistics were calculated in GenAlex v.6.41 (Peakall and Smouse, 2006): the effective number of alleles (A_E) and Shannon's Diversity information index (*I*). To avoid sampling bias, allelic richness (R_S) and private allelic richness (R_{SP}) was computed by implementing rarefaction in HP-RARE1.0 (Kalinowski, 2005). A Kruskal-Wallis test was performed to evaluate significance of differences in the genetic diversity estimates between the generational cohorts.

4.2.4. Effective population size (N_e), Census population size (N_c) and estimation of population bottlenecks

Three methods were used to estimate effective population (N_e) size; the heterozygote excess method, the linkage-disequilibrium method (with minimum allele frequency of 0.02) and temporal method (using F0 as the progenitor for F1; F2 for F3 and F4 for F5) under a random mating LD model in NeEstimator v.1.3 (Peel *et al.*, 2004). To supplement molecular estimates of effective population size, the number of females required to establish the subsequent generation was estimated based on average clutch size for each cohort. Census population size was obtained by

counting all dead flies from each generation (Chapter 2). The ratio of N_e/N_c was also calculated to assess the proportion of total individuals forming the effective population out of the total population. A Wilcoxon signed-rank test for excess heterozygosity (as evidence of a recent population bottleneck) was employed using the infinite alleles model (IAM), the stepwise mutation model (SMM) and the two-phase model (TPM) in Bottleneck v.1.2.02 (Piry *et al.*, 1999). Analysis was performed using 1 000 replications at the 5% nominal level; and TPM was composed of 70% SMM and 30% IAM and a variance of 30.

4.2.5. Relatedness and Inbreeding

Mean within-population pairwise relatedness was computed at the 95% confidence interval (1000 permutations) for each cohort according to the method described by Queller and Goodnight (1989) in GenAlex v.6.41 (Peakall and Smouse, 2006). Inbreeding estimates described by F_{IS}, were also computed in the same software.

4.2.6. Populations differentiation

To evaluate population differentiation estimates, pairwise-F_{ST} was calculated between all cohorts in GenAlex v.6.41 (Peakall and Smouse, 2006). A locus-by locus molecular analysis of variance (AMOVA, 1000 permutations) was computed in the same software. To visualize population distinctness, a Factorial Correspondence Analysis (FCA) was drawn comparing all cohorts in Genetix v.4.05.2 (Belkhir *et al.*, 2004).

4.2.7. Correlations of phenotypic data to genetic diversity statistics

Both Pearson's (r) - and Spearman's (ρ) correlation coefficients were computed at the 5% significance level to test correlations between fitness indicator traits [egg hatchability (%), clutch size (g), pupae weight (g), pupation (%) and female post-mating longevity (% oviposition)] and genetic diversity parameters (H_O, R_S, N_A, F_{IS} and mean relatedness).

4.3. Results

4.3.1. Genetic diversity statistics

Microchecker indicated no genotyping errors due large allele drop out or stuttering, but null alleles occurred in all cohorts (Appendix D: Table D.1), General genetic diversity estimate means across all experimental cohorts are detailed in Table D.1 (Appendix D). Changes of N_A, R_S, R_{SP}, H_O and uH_E across all cohorts are presented in Figure 4.1.

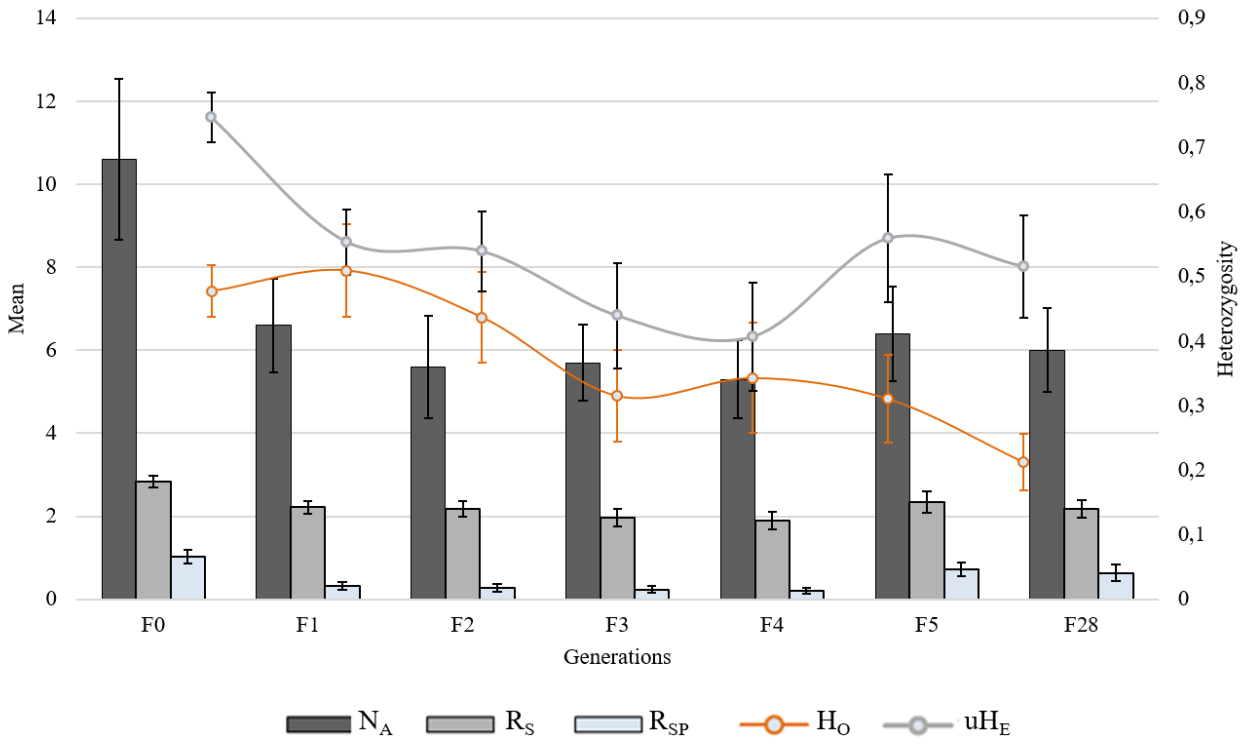

Figure 4.1*:* Basic genetic diversity summary statistic across ten polymorphic microsatellite markers over six experimental generations (F0-F5) in comparison to a commercial population (F28) of *Hermetia illucens*. Presented are the number of alleles (N_A), allelic richness (R_S), private allelic richness (R_{SP}), observed heterozygosity (H_O) and unbiased expected heterozygosity (uH_E).

Significant decreases were observed from F0 to F1 for H_E, uH_E, A_E, I, R_S and R_{SP} ($P \leq 0.01$). A decreasing trend in H_E is seen in F0-F4, while it increases in F4-F5. Observed heterozygosity declines overall, while significant increase in R_{SP} is seen from F4 to F5 ($P = 0.028$). A lower H_O

was apparent in F28, in comparison to the wild founding and experimental populations. Moreover, it was noted that A_E was less than half of N_A in all generations. Average F_{IS} values ranged 0.0469 – 0.0848 for experimental cohorts and all loci deviated significantly from Hardy-Weinberg expectations ($P < 0.005$), showing null alleles across all cohorts (Appendix D: Table D.1).

4.3.2. Effective population size (Ne), Census population size (Nc) and estimation of population bottlenecks

Effective population sizes were generally lower when computed by the temporal method than the linkage disequilibrium method. According to the heterozygosity excess method, N_e sizes were all infinite. The minimum N_e calculated by the linkage disequilibrium method was 7.7 (F3), while the largest was 97 (F5 and F28). Estimates calculated by the temporal method showed N_e can be as small as 3.8 (F0-F1) and as large as 16.9 (F2-F3). For the commercial population, mean N_e was noticeably low as generated from the linkage disequilibrium methods (96.8). Temporal point estimates were similar to the estimated number of females required to produce 0.13g of eggs. For all cohorts N_e/N_c ranged 0.12-0.162 (Table 4.2).

Table 4.1: Effective population size estimates computed by three methods are presented; the heterozygous excess method, the linkage disequilibrium method (random mating model) and the temporal method. Included are also the number of females required to lay 0.13g of eggs, the census population for each cohort (N_c) and the N_e/N_c ratio.

	Estimates						
Method	F0	F1	F2	F3	F4	F5	F28
Heterozygosity excess	∞	∞	∞	∞	∞	∞	∞
Linkage disequilibrium	46.2	78.4	∞	7.7	9.1	96.6	96.8
[parametric 95% CI]	34.1-67.8	34.1-∞	95.4-∞	5.2-10.9	6.1-13.6	33.6-∞	37.3-∞
Temporal		3.8		9.2	3.9		-
[parametric 95% CI]		2.2-6.2		4.9-16.9		2.3-6.3	-
Contributing females	3-32	2-62	2-11	3-17	3-9	*NA	NA
Census population size	2434	484	3517	1938	689	*NA	4000
N_e/N_c	0.012	0.162	0.027	0.004	0.013	*NA	0.024

*Adults did not eclose during F5, therefore clutch weights and N_c could not be obtained.

There was no evidence for significant heterozygous excess assuming any of the mutation models, however under the assumptions of the infinite alleles model (IAM) significant evidence for heterozygote deficiency was seen for F3, while for the two-phase model (TPM; an intermediate between IAM and SMM), significant evidence for heterozygote deficiency was seen for F1, and F28. This indicates population expansion events at these generations and agrees with recordings of census population size (N_c) (Table 4.2).

Table 4.2: Wilcoxon's signed rank test significance values (*P*) for heterozygous deficiency and excess for all cohorts, assuming three mutation models; namely the infinite alleles model (IAM), the two-phase model (TPM) and the stepwise mutation model (SMM). Significant values are denoted in bold.

	Mutation model					
	IAM		TPM		SMM	
	P(He-deficiency)	*P(He-excess)*	*P(He-deficiency)*	*P(He-excess)*	*P(He-deficiency)*	*P(He-excess)*
F0	0.78418	0.24609	0.06543	0.94727	**0.00684**	0.99512
F1	0.09668	0.91992	**0.00342**	0.99756	**0.00244**	0.99854
F2	0.57715	0.46094	0.42285	0.61523	0.08008	0.93457
F3	**0.04199**	0.98389	0.00928	0.99316	**0.00049**	1.00000
F4	0.08008	0.93457	0.00928	0.99316	**0.00244**	0.99854
F5	0.42188	0.62891	0.19141	0.84375	0.00977	0.99414
F28	0.24609	0.78418	**0.00342**	0.99756	**0.00049**	1.00000

4.3.3. Relatedness and Inbreeding

Mean within-population pairwise relatedness estimates are presented in table A4.2 (Appendix A). Individuals in the founding population (F0) were unrelated, relatedness then increases from F0 to F1. From F1, the individuals show increasing relatedness across the experimental period (Figure 4.2).

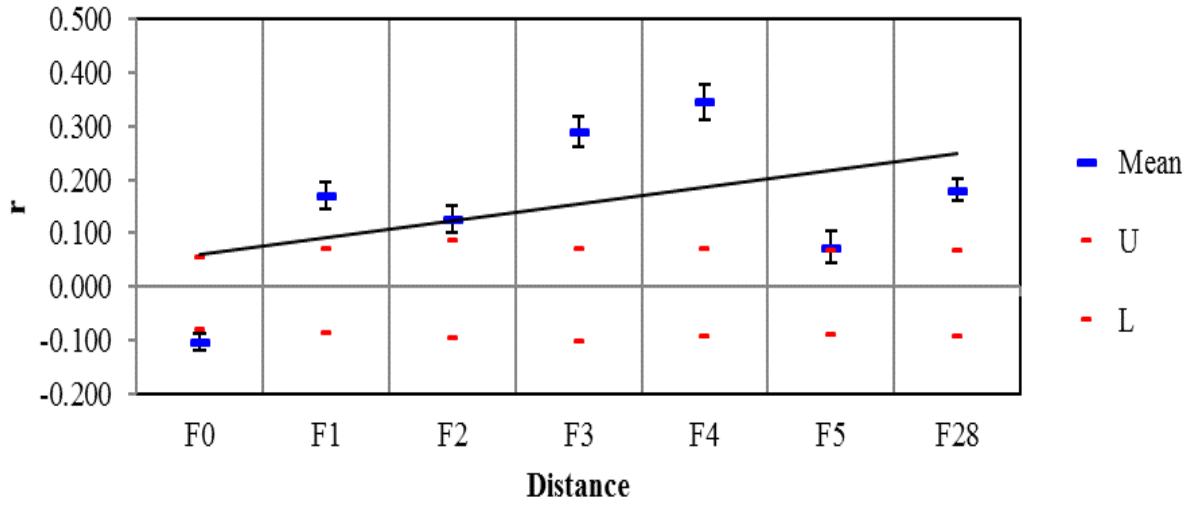

Figure 4.2: Mean within-population pairwise relatedness estimates computed as per Queller and Goodnight (1989). Mean relatedness is demarcated in blue ($y = 0.0314x+0.0293$, $R^2 = 0.2134$); red demarcations indicate upper- and lower confidence intervals (95%).

The commercial population, F28, is also shown in Figure 4.2., as a reference to compare relatedness of experimental cohorts. Only experimental generational cohorts are compared in Figure D.1 and Table D.2 (Appendix D).

4.3.4. Population differentiation

Significant differentiation as described by F_{ST} is seen between all cohorts ($P < 0.01$) and differentiation progressively increases across generational cohorts. The differentiation between F0 and F5 is further emphasized in the FCA plot (Figure 4.3). A lower degree of differentiation between cohorts F1, F2, F3 and F4 is apparent in F_{ST} values. Higher levels of differentiation are seen when comparing F_{ST} values of F5 to all other generations and samples (Table 4.3). These estimates are supported by the clustering of individual genotypes in the factorial correspondent analysis plot (Figure 4.3).

Table 4.3: Pairwise F_{ST} values as calculated in GenAlex v.6.41 for all generational cohorts, including a commercial population (F28) bottom diagonal. Significance values are depicted top diagonal.

	F0	F1	F2	F3	F4	F5	F28
F0	-	0.001	0.001	0.001	0.001	0.001	0.001
F1	0.099^{B}	-	0.001	0.001	0.001	0.001	0.001
F2	0.060^{B}	0.036^{A}	-	0.001	0.001	0.001	0.001
F3	0.118^{B}	0.130^{B}	0.065^{B}	-	0.002	0.001	0.001
F4	0.127^{B}	0.135^{B}	0.064^{B}	0.031^{A}	-	0.001	0.001
F5	0.186^{C}	0.282^{D}	0.271^{D}	0.343^{D}	0.306^{D}	-	0.001
F28	0.097^{B}	0.102^{B}	0.087^{B}	0.138^{B}	0.168^{C}	0.291^{D}	-

A = little genetic differentiation
B = moderate genetic differentiation
C = great genetic differentiation
D = very great genetic differentiation

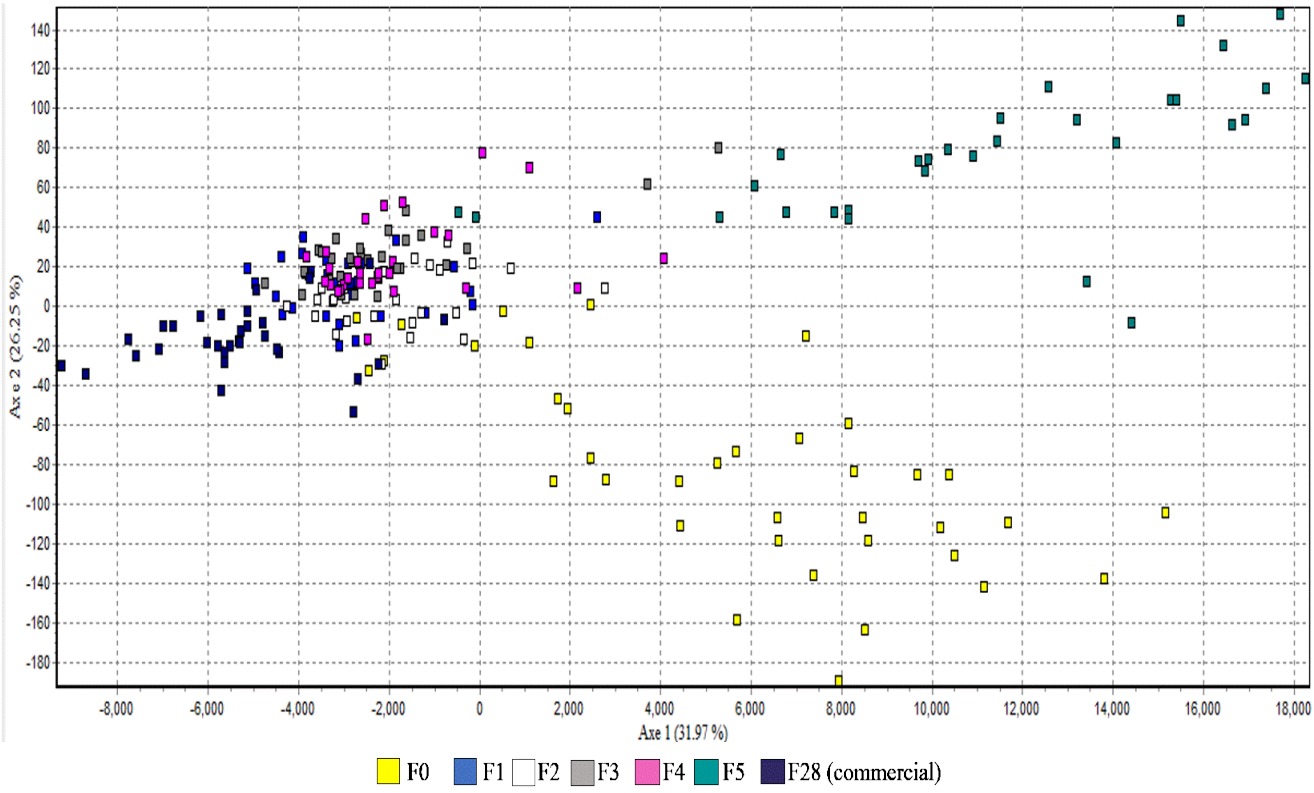

Figure 4.3: Factorial correspondence analysis plot showing individual genotype data for all experimental populations (F0-F5), as well as a commercial population (F28).

Population differentiation is also supported by AMOVA results, with genetic differentiation amongst generations (F_{IS} = 0.369, $P < 0.00$); within generations (F_{IT} = 0.469; $P < 0.00$) and over all generations (F_{ST} = 0.158; $P < 0.00$); with most variance lying within individuals (Table 4.4).

Table 4.4: Molecular analysis of variance results base on ten loci for all cohorts.

Source of variance	Sum of Squares	Variance components	Variation (%)
Amongst Generations	214.262	0.534	16
Amongst Individuals	785.958	1.050	31
Within Individuals	374.500	1.792	53
Total	1374.720	3.375	100%

4.3.5. Correlations of phenotypic data to genetic diversity statistics

Pearson's and Spearman's correlation coefficients and significance values are presented in Table D.3 (Appendix D), but the most noteworthy were pupation and clutch weight exhibiting a significant negative correlation with H_O (r = -0.9808, P = 0.003 and r = -0.9123, P = 0.03, respectively) (Figure 4.4 a and b). Furthermore, a significant postive correlation is exists between H_O and % oviposition (r = 0.9253, P = 0.02) (Figure 4.4 c).

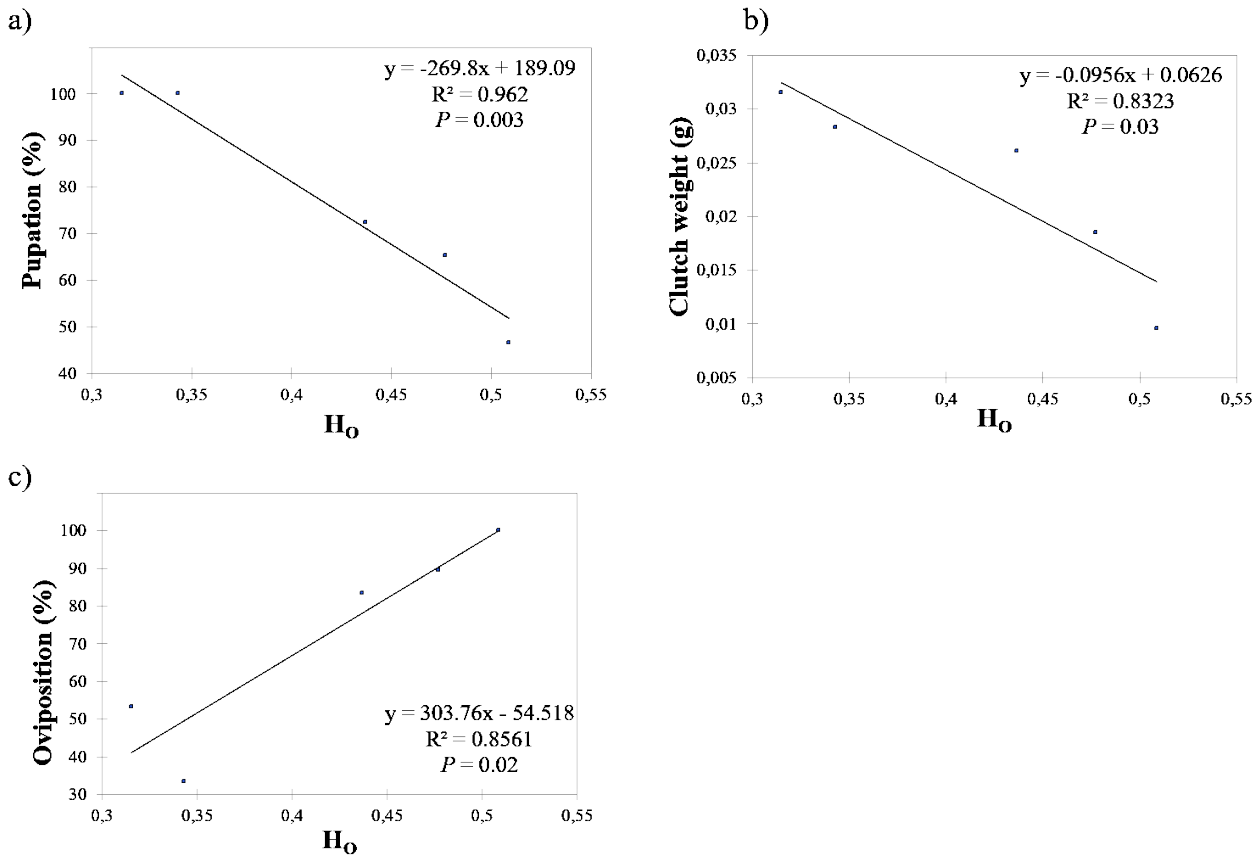

Figure 4.4: Observed heterozygosity (H_O) was correlated to fitness-related traits. Presented here are % pupation (a), clutch weight (b) and % oviposition (c) by implementing Pearson's correlation coefficient (*r*). Significance of correlations are represented by *P*. Indicated in the figure is the regression value (R^2).

Pupae weight exhibits a significant negative correlation between N_A and R_S (r = -0.9036, P = 0.03 and r = -0.8953; P = 0.04, respectively) (Figure 4.5).

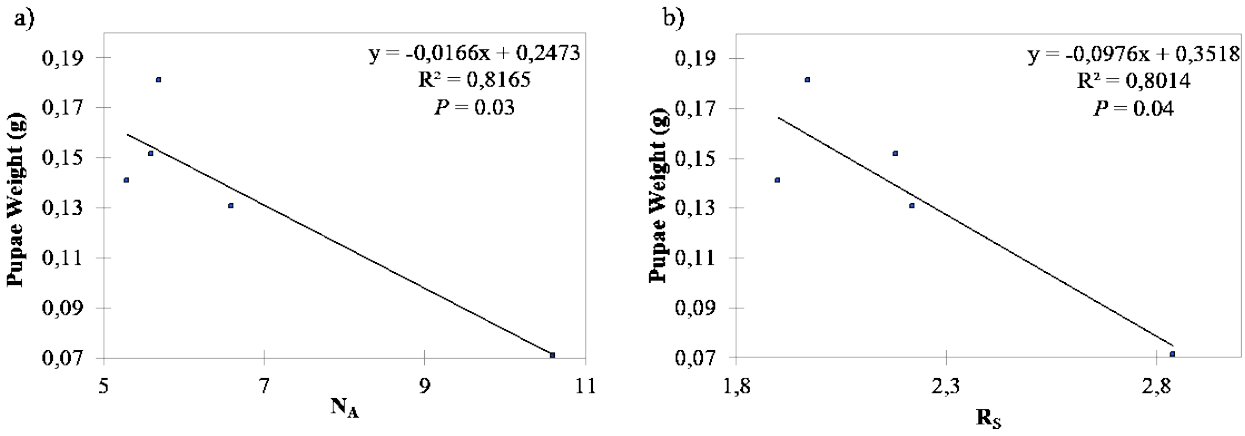

Figure 4.5: Number of alleles (N_A) (a) and allelic richness (R_S) (b) were correlated to an indirect indicator of adult size and fecundity (*i.e.* pupae weight) by implementing Pearson's correlations (*r*), significance of correlations are represented by *P*. Indicated in the figure is the regression value (R^2).

4.4. Discussion

4.4.1. Genetic diversity, effective population, bottlenecks and relatedness

Within all cohorts, significant violation of Hardy-Weinberg assumptions was noted. This was mainly due to homozygous excess, based on positive F_{IS} estimates. Homozygous excess could have been the result of null alleles, which were detected in high frequency in all cohorts and occur commonly in insects (Chapuis and Estoup, 2006), or as a result of inbreeding.

Results indicated a significant loss of genetic diversity from the wild founding population (F0) to the final generation (F5) based on estimates such as number of alleles, allelic richness and heterozygosity. This is to be expected and has been well documented in previous studies focussed on establishing artificial colonies from the wild (Mayr, 1954; Miyatake 1998; Baek *et al.*, 2014). Furthermore, low levels of genetic diversity were observed for the commercial population. This is unsurprising as the colony is four years old and has never been supplemented with new individuals since the founding event. This is a strong indication of the implications for not accounting for genetic diversity during colony establishment and throughout colony management; and emphasises the importance of assessing genetic diversity. It is noted that the commercial population maintains

an extremely low effective population size as well, which would further explain the low levels of genetic diversity (Table 4.2).

There was a large discrepancy between the number of alleles and the number of effective alleles in all generations. Alleles occurring at low frequencies in the founding population are easily lost by random genetic drift. Furthermore, a significant increase in private allelic richness was observed in the final generation. This may be additional evidence for genetic drift as private alleles are unmasked. However, the commercial population also showed large private allelic richness. During the larval rearing stages of experiments described in Chapter 2, there was an instance where the layer hen feed was left uncovered. Escaped females form the commercial population may have oviposited in the feed, resulting in cross - contamination of strains. This may also explain the increase in heterozygosity in F5 in comparison to preceding generations.

Gene diversity is expected to affect short-term response to selection, but allelic richness may affect long-term responses to selection, and eventually population persistence (Allendorf, 1986). This study did not directly assess selection, but emphasises the importance of screening initial genetic diversity of founding populations and further monitoring fluctuations throughout mass-rearing. The loss of alleles can be slowed down by maintaining a large effective population size (Forstmeier *et al.,* 2007). The LD estimate of effective population size provides a better estimate of effective population size (Waples and Do, 2010). Mean effective population estimates based on the LD method were substantially low. In addition, the ratios of N_e/N_c were well below what is required for sustained genetic diversity (Hedrick, 2005). In the present study, N_e/N_c ratios ranged 0.01 – 0.16; while it is recommended that they range 0.25-0.75. This further explains the substantial decline in genetic diversity observed throughout this study. Significant homozygous excess was evident using Wilcoxon's test for heterozygous deficiency; this indicates population expansion and agrees with increased census population size at the respective generations where significance of heterozygous deficiency was detected. Cumulatively, results present evidence for population bottlenecks brought

on by low effective population size, which results in increased homozygosity. The strategy implemented to establish successive generations lends itself to significantly decreasing the effective contribution of individuals from preceding generations, and this is illustrated by the decreased effective population at generations where clutch sizes are larger. More females are required to contribute to the required 0.13g of eggs (Chapter 2) when they produce smaller clutch sizes. This strategy of establishing successive generations may also explain why the effective population size of the commercial colony is low, and also provides evidence that random genetic drift is responsible for the occurrence of private alleles in F5 and F28.

In addition, the decrease in census population size seen at different generations (Table 4.2) would increase the probability of inbreeding. The likelihood of inbreeding is further confirmed by the increased estimates of relatedness across generations (Figure 4.2) – further explaining the loss of overall genetic diversity. A decline in overall heterozygosity and allelic richness has been experienced in other studies where mean relatedness also gradually increases across generations, similar to this study (Willoughby *et al.*, 2015).

4.4.2. Population differentiation

Pairwise F_{ST} estimates were significant amongst all generations and indicated a gradual increased differentiation toward the end of the study period (Table 4.3). Moreover, genetic differentiation noted between the wild, F5 and commercial populations agrees with the three distinct separations obtained with the factorial correspondence plot (Figure 4.3). There is further evidence for population differentiation within each cohort as demonstrated by the AMOVA and pairwise F_{ST} results. Pairwise F_{ST} estimates between F5 and the commercial colony were less pronounced than for instance F5 and F4. This may also be evidence of cross-contamination events described earlier.

Santos *et al.* (2012) also observed abrupt increase in the genetic differentiation in *Drosophila subobscura* within first generations of colony establishment. The differentiation observed for the

experimental generations may be indicative of a slow adaptation to artificial environments. The markers used in this investigation were not suited for analyses on selection. However, in another study where intentional artificial selection pressures were tested, the genetic divergence of *G. tokunagai* also increased rapidly in the early generations, while stabilizing at the 11th generation (Nowak *et al.*, 2007)

4.4.3. Correlations of phenotypic data to genetic diversity statistics

A significant positive relationship between post-mating longevity (described as the number of females which survive until oviposition, (% oviposition) and observed heterozygosity is found in this study. The decline in adult longevity, relative to heterozygosity has been well established. These results confirm the effects of a loss of heterozygosity on female fitness- an indication of inbreeding depression (Charlesworth and Charlesworth, 1987; Charlesworth and Willis, 2009). It was also demonstrated that male longevity was not affected by inbreeding in a study by Messina *et al.* (2013). Male longevity was not recorded in the present study, but could explain the decrease in female fecundity measurements (Chapter 2). Asynchronous survival between sexes would also result in the decreased frequency of mating and therefore egg production. Messina *et al.* (2013) also found that inbred females adapted their oviposition behaviour to favour an experimental design chosen to maximize egg production. However, this did not improve fecundity in *Callosbruchus chinensis*. This also agrees with the correlation between clutch weight and observed heterozygosity in this study.

A significantly positive correlation between relatedness of individuals in a cohort and pupae weight occurred, whereas there was a significantly negative correlation between number of alleles (N_A) and allelic richness (R_S) and pupae weight. In holometabolic insect mass-rearing systems, pupae are stocked to attain a predetermined adult density in the cages, based on bulk pupae - weight. This strategy would result in decreases of population density as pupae weights increase. Essentially, when pupae weights are higher, fewer pupae would be stocked in a cage - resulting in fewer adults

and a lower effective population size. Since pupae stocking is also synchronised in mass-rearing systems, it is likely that pupae of the same age (related individual) would mature at the same time; this is evident in the correlation between pupation and observed heterozygosity. In combination, results indicate a continuous reduction in genetic diversity by the pupae raffling process (Gilchrist *et al.*, 2012).

4.5. Conclusion

The resulting decrease in genetic diversity caused by the decrease in effective populations size (*i.e.* the founder effect) is well illustrated in this chapter. Overall the rapid reduction genetic diversity was correlated to a number of commercially relevant traits and provides evidence of the deleterious effects of inbreeding depression and genetic drift. This was affected by the experimental deign and ultimately reflects the impact of general mass-rearing strategies pertaining to holometabolous insects. Ultimately, the vulnerability of genetic diversity throughout colony establishment is demonstrated. This study has implications for colony management strategies and can aid in optimising them to conserve genetic diversity. Care should be taken when managing the initial genetic diversity of founding members. Specific attention should be paid to facilitating mating success and ensuring an increase in effective population size.

Chapter 5: Concluding Remarks

5.1. Research overview

The growing interest in exploiting *Hermetia illucens* as a waste-recycling solution has driven research relevant to the biology and life-history of this fly. However, specific investigations into the domestication and mass-rearing of *H. illucens* have thus far been overlooked. Understanding the decline in genetic diversity associated with poor colony management is invaluable as it affects commercially-relevant phenotypes. The aim of this study was to gain an understanding of the bottlenecks occurring during establishing commercial colonies of *H. illucens*. The "longitudinal approach" described by Price (2002) was implemented, and both phenotypic and genetic sampling was conducted at each of six successive generational cohorts. The tools required to investigate genetic diversity were developed in in the form of microsatellite markers. This study ultimately gave insight in to the different genetic mechanisms at play during colony establishment and how they are detrimental to captive-bred populations.

Changes in life history characteristics were experienced while attempting to establish a colony from wild individuals (Chapter 2). To supplement the investigations into the genetic mechanisms related to these changes, a panel of ten polymorphic microsatellite markers was developed and characterised (Chapter 3). The markers were compiled in to three multiplex PCR reactions to facilitate cost-effective data collection. This novel set of markers was able to efficiently detect genetic diversity in *H. illucens* and was further used in the assessment of genetic diversity during colony establishment (Chapter 4). Aside from gaining an understanding of the phenotypic trends observed in Chapter 2, Chapter 4 also gave insight in to the demographic effects on genetic diversity.

5.2. Population dynamics under commercial conditions

Initial phenotypic measurements such as clutch weight and pupae weight increased from F0 as the flies acclimated to the artificial environment. This is typical of newly isolated populations, where reduction in population size (N_c) is also experienced, while the organism responds to new environmental conditions (Shimoji and Miyatake 2002; Baeshen *et al.*, 2014; Zygouridis *et al.*, 2014). Overall, recordings of phenotypic variables declined toward the end of the experimental period. Notable decreases in pupae weight, post-mating longevity (% oviposition), clutch size and hatchability occurred toward the end of the experimental period. The size of the pupa determines the size of the adult, and smaller females will have less energy stored in their fat bodies to survive until oviposition. Oogenesis is an energy-limited process, and a limited amount of stored energy would result in the reabsorbing of oocytes to conserve energy, decreasing clutch size. Moreover, a decrease in female size has also been reported to decrease egg hatchability, but in this study a correlation between hatchability and pupae weight was not evident (Furman *et al.*,1959; Davidowitz *et al.*, 2003; Pastor *et al.*, 2011). Population size (N_c) fluctuated dramatically across generations and this is also attributed to the stochastic nature of BSF hatchability and it's correlation to relative humidity (which showed variability across cohorts). Furthermore, an exceptionally low N_e/N ratio was calculated across all generational cohorts. This decreased overall genetic diversity significantly across all cohorts and contributed to the increase population homozygosity. The impact of genetic diversity on relevant phenotypes during early stages of colony establishment was well illustrated in this study. Most notable were the significant correlations between observed heterozygosity and pupation (%), clutch weight (g) and post-mating longevity (% oviposition); as well as pupae weight and number of alleles and allelic richness.

Clutch weight and pupation appear to be correlated variables (Chapter 2). The decrease in clutch weight would result in the sampling of a larger variety of clutches for the next generation. In generations where clutch weights decrease, N_e increases. It may be for this reason that H_O was

significantly negatively correlated to clutch weight and pupation (%). Indirectly, this alludes to a possible strategy of increasing H_O in the mass-rearing system. By intentionally selecting clutches from a wide range of individuals, N_e and H_O can be increased. In addition, the significant relationship between decreased H_O and decreased post-mating longevity (% oviposition) gives evidence of inbreeding depression. This agrees with a study by Messina *et al.* (2013), where the crossing of related individuals resulted in decreased female longevity. Furthermore, as genetic diversity, described by N_A and R_S, decreases, so does pupae weight. Pupae weights are generally an indirect measure of adult fitness as previously described, further indicating how the loss of genetic diversity can impact the fitness of the colony (*i.e.* evident of inbreeding depression). Previous studies have shown similar results and this serves as a cautioning to the effects of a loss of heterozygosity (Fox *et al.*, 2006).

Previously, the lack of understanding of which commercially-relevant traits are negatively impacted by early domestication led to lack of understanding of how these changes will translate into decreased colony performance. This effect was predicted in other studies where commercially-relevant traits suffered during the domestication process, but is now illustrated in Chapter 4, as well (Ferguson *et al.*, 2005; Howell and Knols, 2009).

5.3. Management considerations

The origin and size of the founding population plays an important role in allowing the greatest amount of variation and plasticity during domestication. The failure to sustain a colony adapted to the artificial rearing environment as seen in F5, is not an uncommon occurrence (Chapter 2). For example, Baek *et al.* (2014) failed to establish artificial colonies of *Glyptotendipes tokunagai* further than the 5th generation, similar to this study. Once the founding population was increased four-fold – thereby increasing genetic heterogeneity - Baek *et al.* (2014) managed to sustain a colony beyond 22 generations. Developmental time and waste conversion varies across populations

of *H. illucens* and should therefore be considered when selecting and establishing a population as a waste management agent in different regions of the world (Zhou *et al.*, 2013). This may also be advantageous when considering selective breeding strategies, which are discussed later. Moreover, wild individuals caught in areas with vastly fluctuating temperatures may struggle to adapt and equilibrate to constant temperate temperatures (Brakefield and Kesbeke, 1997). The details of this process differ across species depending on factors such as the periodicity of feeding and food assimilation, and specific attention to should be paid to the species of interest, rather than generalising effects.

By implementing a strategy typical of insect mass-rearing, this study aided expanding this knowledge and makes it possible to improve on current breeding schemes and mass-rearing practices. The production strategy outlined in Chapter 1 lends itself to a variety of inadvertent selection pressures and risks associated with stochastity in the system. Typically, offspring are combined and shuffled when larvae originating from different cages are reared in one collective trough. However, the N_e contributing to the next stocking event varies. For example, if the required amount of eggs is obtained from a single cage, that single cage contributes to the new population. If 50 cages are required to make up the next batch of larvae, the effective population and variation is increased. Moreover, the next series of cages are stocked from a series of pupae collections from different harvesting days, if this is done from different cages every day, more variation is added to the next stocking series. This entire system is affected by the phenotypic parameters assessed in this study. To ensure the quality and efficacy of the process, it is vital to ensure the health and fecundity of the adults and robustness of immature stages. Quality control parameters are regularly recorded; however, they must be applied to adjust stocking ratios of pupae and variability in N_e by selecting clutches from a wide range of females. This study further highlights the importance of considering N_e, rather than depending solely on census population size, which is a common drawback in some of

the holometabolic insect mass-rearing systems described (Sheppard *et al.,* 2002; Ekesi, 2007; Vantomme *et al.,* 2012).

The most common suggestion to counteract the effects of inbreeding in mass-reared colonies is augmentation of the effective population. This approach was successful in *Bacterocera olaea* colonies. Practical solutions to managing the inevitable loss of genetic diversity are vital to any mass-rearing system, as strategies must be implemented on a large scale. The anticipated negative effects of inbreeding on laboratory-reared mosquitoes, for example, have led to different strategies for re-forming their genetic diversity prior to mosquito release programmes (Benedict *et al.,* 2009). In these approaches, crossing and backcrossing laboratory strains with the progeny of field-collected individuals is required; this is not always practical to implement regularly and efficiently and is considered haphazard (Benedict *et al.,* 2009).

5.4. Shortcomings of project and future undertakings

One clear limitation to this study is that only a single population was assessed. In addition to this, the founding members of this population came from the same sampling site. An expansion on this study could involve establishing multiple populations and sustaining them simultaneously. This would expand on the understanding of how isolated populations from the same source may differentiate independently. A further expansion of this could be to investigate the exploitation of hybrid vigour, and assess the phenotypic changes associated with this approach.

The impact of genetic management of captive colonies is improved by exploiting population genetics. Understanding the levels of genetic diversity and population demographics improves the design of strategies to manage captive populations. There are a variety of molecular markers and population genetic estimates which assist in these designing these strategies. In this study, only ten microsatellite markers were used. Although they were sufficient in giving insight into the assessment of genetic diversity in this study, there is always the possibility of increasing resolution

and expanding research by developing more markers. These could include single nucleotide polymorphisms (SNP's); expressed sequenced tags (EST's) and mitochondrial markers.

A difference is seen in both temporally and geographically different samples due to shifts in the genetic composition of populations. This affects the repeatability of experiments investigating laboratory evolution, as well as comparisons of different life-history traits drawn from other studies regarding *H. illucens* biology (Simões *et al.*, 2008). Understanding fundamental differences in strains may have implications globally, such data will also allow for more informed decisions when selecting a strain for initiating a new facility (Zhou *et al.*, 2013). Therefore, it is recommended that a different set of markers be developed to study the population structure of wild individuals for future colony supplementation. These could include mitochondrial markers, which would give insight in to the ancestral global genetic structure of *H. illucens*. *Hermetia illucens* is thought to have been introduced to KwaZulu-Natal, South Africa from the New World (between 1915 and 1945). Aside from labelled specimens, no databases exist in which the global *H. illucens* population is listed (Üstüner *et al.*, 2003, Roháček *et al.*, 2013; Marshall *et al.*, 2015). Investigating the global genetic resources of *H. illucens* and applying a population genomics approach can improve selective breeding protocols. Genomic analysis of diverse populations can be used to uncover the genetic basis of complex traits. For example, genome-wide single nucleotide polymorphism scans can identify loci under selection and dissect quantitative traits, thereby elucidating the genetic basis of commercially-relevant traits. This information can be used to design selective breeding strategies targeting specific traits. Another prospective research endeavour could be to implement a marker-assisted selective (MAS) breeding scheme. Combing GWAS data and implementing MAS has been done previously on other insects, such as the domesticated silkworm, *Bombyx mori*, where silk production was successfully increased (Nagaraju, 2002). As reports by Nagaraju (2002) also suggest, an extensive knowledge of the organism's genome is required to successfully implement MAS. Aside from whole genome shotgun sequences (represented by one male and one female)

available on the NCBI Genbank, no other sequence data on BSF is available (Bachtrog and Vicoso, 2015).

5.5. Final Remarks

As an upcoming industry, the approaches to BSF mass-rearing can be significantly improved. The interplay of phenotypic, genetic and environmental interactions is displayed in this study. Importantly, all these factors should be considered not only during initial stages of colony establishment, but throughout monitoring and optimising the commercial mass-rearing process. From this study, it is clear that the control of environmental conditions, such as humidity and lekking space should be strictly managed to optimize the mass-rearing of *H. illucens* (*i.e.* the effect of humidity correlated to hatchability significantly, while genetic diversity estimates did not**)**. It is also vital to understand intrinsic processes which may lead to a decline of production efficiency, but which can in turn be manipulated in favour of mass-rearing. These processes included how changes of one phenotype in a preceding generation may affect another phenotype in a successive generation, as seen in the effect of pupae size on egg clutch size (Chapter 2). Moreover, the value of using molecular markers and population genetics statistics is illustrated in this study. There is great potential for expanding on this study, which has already yielded valuable insight. The interplay between genetics and phenotypes was very evident. The impact of population bottlenecks during the founding event was demonstrated, in addition to the compounded effect on genetic diversity when colony management does not accommodate phenotypic changes (such as increased pupae weight). Overall, the results obtained in this study can aid in establishing new colonies and improving strategies for management of already existing colonies of *Hermetia illucens*.

References

AARNES, S. G., FLØYSTAD, I., SCHREGEL, J., VINDSTAD, O. P. L., JEPSEN, J. U., EIKEN, H. G., and HAGEN, S. B. (2015). Identification and evaluation of 21 novel microsatellite markers from the autumnal moth (*Epirrita autumnata*) (Lepidoptera: Geometridae). *International journal of molecular sciences*, 16(9), 22541-22554.

AKETARAWONG, N., CHINVINIJKUL, S., ORANKANOK, W., GUGLIELMINO, C. R., FRANZ, G., MALACRIDA, A. R., and THANAPHUM, S. (2011). The utility of microsatellite DNA markers for the evaluation of area-wide integrated pest management using SIT for the fruit fly, *Bactrocera dorsalis* (Hendel), control programs in Thailand. *Genetica*, 139(1), 129-140.

ALCOCK, J. (1990). A large male competitive advantage in a lekking fly, *Hermetia comstocki* Williston (Diptera: Stratiomyidae). *Psyche*, 97(3-4), 267-280.

ALDER, J., CAMPBELL, B., KARPOUZI, V., KASCHNER, K., and PAULY, D. (2008). Forage fish: from ecosystems to markets. *Annual review of environment and resources*, 33(1), 153.

ALLENDORF, F.W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, 5, 181–190.

AMOS, W., HOFFMAN J.I., FRODSHAM A., ZHANG L., BEST S. and HILL V.S. (2007). Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes*, 7, 10-14.

ANNE, C. (2006). Choosing the right molecular genetic markers for studying biodiversity: from molecular evolution to practical aspects. *Genetica*, 127(1-3), 101-120.

BACHTROG, D., WEISS, S., ZANGERL, B., BREM, G., and SCHLÖTTERER, C. (1999). Distribution of dinucleotide microsatellites in the *Drosophila melanogaster* genome. *Molecular Biology and Evolution*, 16(5), 602-610.

BAEK, M. J., YOON, T. J., KANG, H. J., and BAE, Y. J. (2014). Biological and Genetic Characteristics of *Glyptotendipes tokunagai* (Diptera: Chironomidae) on the basis of successive rearing of forty-two generations over seven years under laboratory conditions. *Environmental Entomology*, 43(5), 1406-1418.

BAESHEN, R., EKECHUKWU, N. E., TOURE, M., PATON, D., COULIBALY, M., TRAORÉ, S. F., and TRIPET, F. (2014). Differential effects of inbreeding and selection on male reproductive phenotype associated with the colonization and laboratory maintenance of *Anopheles gambiae*. *Malaria journal*, 13(1), 1-14.

BANKS, I. J., GIBSON, W. T., and CAMERON, M. M. (2014). Growth rates of black soldier fly larvae fed on fresh human faeces and their implication for improving sanitation. *Tropical Medicine and International Health*, 19(1), 14-22.

BARROS-CORDEIRO, K. B., BÁO, S. N., and PUJOL-LUZ, J. R. (2014). Intra-puparial development of the black soldier-fly, *Hermetia illucens*. *Journal of Insect Science*, 14(1), 83.

BARTLETT, A.C. (1993). Maintaining genetic diversity in laboratory colonies of parasites and predators. In: Narang, S.K., Bartlett, A.C. and Faust, R.M. (eds) *Applications of Genetics to Arthropods of Biological Control Significance*: CRC Press, Boca Raton, Florida, pp. 133–145.

BEILHARZ, R. G., LUXFORD, B. G., and WILKINSON, J. L. (1993). Quantitative genetics and evolution: is our understanding of genetics sufficient to explain evolution? *Journal of Animal Breeding and Genetics*, 110 (1-6), 161-170.

BELKHIR, K., BORSA, P., CHIKHI, L., RAUFASTE, N., and BONHOMME, F. (1996). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. *Laboratoire génome, populations, interactions*, CNRS UMR, 5000, 1996-2004.

BENEDICT, M. Q., KNOLS, B. G., BOSSIN, H. C., HOWELL, P. I., MIALHE, E., CACERES, C., and ROBINSON, A. S. (2009). Colonisation and mass rearing: learning from others. *Malaria Journal*, 8(2), S4.

BERG, I., NEUMANN, R., CEDERBERG, H., RANNUG, U. AND JEFFREYS, A. J. (2003). Two modes of germline instability at human minisatellite MS1 (locus D1S7): complex rearrangements and paradoxical hyperdeletion. *The American Journal of Human Genetics*, 72, 1436–1447.

BILDE. T, MAKLAKOV A, MEISNER K, LA GUARDIA L, FRIBERG U (2009). Sex differences in the genetic architecture of lifespan in a seed beetle: extreme inbreeding extends male lifespan. *BMC Evolutionary Biology*, 9: 33.

BODRI, M. S., and COLE, E. R. (2007). Black soldier fly (*Hermetia illucens* Linnaeus) as feed for the American Alligator (*Alligator mississippiensis* Daudin). *Georgia Journal of Science*, 65(2), 82-88.

BONDARI, K. AND SHEPPARD, D.C., (1981). Soldier fly larvae as feed in commercial fish production. *Aquaculture*, 24: 103-109.

BONDARI, K., and SHEPPARD, D. C. (1987). Soldier fly, *Hermetia illucens* L., larvae as feed for channel catfish, *Ictalurus punctatus* (Rafinesque), and blue tilapia, *Oreochromis aureus* (Steindachner). *Aquaculture Research*, 18(3), 209-220.

BOOTH, D. C., and SHEPPARD, C. (1984). Oviposition of the black soldier fly, *Hermetia illucens* (Diptera: Stratiomyidae): eggs, masses, timing, and site characteristics. *Environmental Entomology*, 13(2), 421-423.

BRADLEY, G. H. (1930). *Hermetia illucens* L. A Pest of sanitary Privies in Louisiana. *Journal of Economic Entomology*, 23(6), 1012-1013.

BRADLEY, S. W., and SHEPPARD, D. C. (1984). House fly oviposition inhibition by larvae of *Hermetia illucens*, the black soldier fly. *Journal of chemical ecology*,10(6), 853-859.

BRAKEFIELD, P. M., and KESBEKE, F. (1997). Genotype–environment interactions for insect growth in constant and fluctuating temperature regimes. *Proceedings of the Royal Society of London B: Biological Sciences*, 264(1382), 717-723.

BRENT, C. S. (2010). Stage-specific effects of population density on the development and fertility of the western tarnished plant bug, *Lygus hesperus*. *Journal of Insect Science*, 10(49), 1-15.

BROOKFIELD, J. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, 5(3), 453-455.

CALVO, D., and MOLINA, J. M. (2005). Fecundity–body size relationship and other reproductive aspects of *Streblote panda* (Lepidoptera: Lasiocampidae).*Annals of the Entomological Society of America*, 98(2), 191-196.

CASTAÑÉ, C., ALOMAR, Ò., RIUDAVETS, J., and GEMENO, C. (2007). Reproductive biology of the predator *Macrolophus caliginosus*: effect of age on sexual maturation and mating. *Biological Control*, 43(3), 278-286.

CASTAÑÉ, C., BUENO, V. H., CARVALHO, L. M., and VAN LENTEREN, J. C. (2014). Effects of founder population size on the performance of *Orius laevigatus* (Hemiptera: Anthocoridae) colonies. *Biological Control*, 69, 107-112.

CHAKRABORTY, R., DE ANDRADE, M., DAIGER, S.P., BUDOWLE, B. (1992). Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of Human Genetics*, 56, 45-47.

CHAMBERS, D. L. (1977). Quality control in mass rearing. *Annual Review of Entomology*, 22(1), 289-308.

CHAMBERS, G. K., and MACAVOY, E. S. (2000). Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 126(4), 455-476.

CHAPMAN, J. R., NAKAGAWA, S., COLTMAN, D. W., SLATE, J., and SHELDON, B. C. (2009). A quantitative review of heterozygosity–fitness correlations in animal populations. *Molecular Ecology*, 18(13), 2746-2765.

CHAPUIS, M.P., ESTOUP, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, 24(3), 621– 631.

CHARLESWORTH, D. and CHARLESWORTH, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*, 18, 237–268.

CHARLESWORTH, D., and WILLIS, J. H. (2009). The genetics of inbreeding depression. *Nature Reviews Genetics*, 10(11), 783-796.

CHARLTON, A. J., DICKINSON, M., WAKEFIELD, M. E., FITCHES, E., KENIS, M., HAN, R.,and BRUGGEMAN, G. (2015). Exploring the chemical safety of fly larvae as a source of protein for animal feed. *Journal of Insects as Food and Feed*, 1(1), 7-16.

CHOI, Y. C., LEE, S. B., PARK, K. H., NAM, S. H., KIM, J. G., KIM, W. T., and CHOI, J. Y. (2012). Artificial Multiplication of the Black Soldier Fly (BSF), *Hermetia illucens* (Diptera: Stratmyidae) Using the Livestock Feces. *Journal of Sericultural and Entomological* Science, 50(2), 63-70.

CHRISTIAKOV, D. A., HELLEMANS, B., VOLCKAERT, F.A.M. (2006). Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture*, 255, 1-29.

CHU, N.D., KALUZIAK, S. T., TRUSSELL, G. C., AND VOLLMER, S. V. (2014). Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail *Nucella lapillus. Molecular Ecology*. 23, 1863–1873.

ČIČKOVÁ, H., NEWTON, G. L., LACY, R. C., and KOZÁNEK, M. (2015). The use of fly larvae for organic waste treatment. *Waste Management*, 35, 68-80.

COLLARD, B. C. Y., JAHUFER, M. Z. Z., BROUWER, J. B., and PANG, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytic*a, 142(1-2), 169-196.

COOPER, G., (1995). Analysis of genetic variation and sperm competition in dragonflies. D. Phil. Thesis, Oxford University.

COPELLO, A. (1926). Biologia de *Hermetia illucens* L. *Revista de la Sociedad Entomológica Argentina*, 1(2), 23-7.

CRNOKRAK, P., and ROFF, D. A. (1999). Inbreeding depression in the wild. *Heredity*, 83(3), 260-270.

CROW, J. F., and KIMURA, M. (1970). An introduction to population genetics theory. *An introduction to population genetics theory*. Harper and Row, New York.

DAKIN EE, AVISE JC (2004) Microsatellite null alleles in parentage analysis. *Heredity*, 93: 504-509.

DAVIDOWITZ, G., D'AMICO, L. J., and NIJHOUT, H. F. (2003). Critical weight in the development of insect body size. *Evolution and development*, 5(2), 188-197.

DEKA R., GUANGYUN S., SMELSER D., AHONG Y., KIMMEL M. AND CHAKRABORTY R. (1999). Rate and directionality of mutations and effects of allele size constraints at anonymous, gene-associated and disease-causing trinucleotide loci. *Molecular Biology and Evolution*, 16(9)1166-1177.

DI MAIO, A., and DE CASTRO, O. (2013). SSR-patchwork: an optimized protocol to obtain a rapid and inexpensive SSR library using first-generation sequencing technology. *Applications in plant sciences*, 1(1), 1200158.

DIENER, S., SOLANO, N. M. S., GUTIÉRREZ, F. R., ZURBRÜGG, C., and TOCKNER, K. (2011). Biological treatment of municipal organic waste using black soldier fly larvae. *Waste and Biomass Valorization*, 2(4), 357-363.

DIENER, S., ZURBRÜGG, C., and TOCKNERA, K. (2009). Conversion of organic material by black soldier fly larvae – Establishing optimal feeding rates. *Waste Management and Research*.27(6), 603-610

DIERENFELD, E. S., and KING, J. D. (2008). Digestibility and mineral availability of Phoenix worms, *Hermetia illucens*, ingested by mountain chicken frogs, *Leptodactylus fallax*. *Journal of Herpetological Medicine and Surgery*, 18(3/4), 100-105.

DIXON, T. J., COMAN, G. J., ARNOLD, S. J., SELLARS, M. J., LYONS, R. E., DIERENS, L., and LI, Y. (2008). Shifts in genetic diversity during domestication of Black Tiger shrimp, *Penaeus monodon*, monitored using two multiplexed microsatellite systems. *Aquaculture*, 283(1), 1-6.

DOBNEY, K., and LARSON, G. (2006). Genetics and animal domestication: new windows on an elusive process. *Journal of Zoology*, 269(2), 261-271.

DRAYTON, J. M., and JENNIONS, M. D. (2011). Inbreeding and measures of immune function in the cricket *Teleogryllus commodus*. *Behavioral Ecology*, *22*(3), 486-492.

DREW, D. J. W., and PIETERSE, E. (2015). Markets, money and maggots. *Journal of Insects as Food and Feed*, 1(3), 227-231.

ECONOMOPOULOS, A. P., and LOUKAS, M. (1986). ADH allele frequency changes in olive fruit flies shift from olives to artificial larval food and *vice versa*, effect of temperature. *Entomologia experimentalis et applicata*, 40(3), 215-221.

EISEN, J. A. (1999). Mechanistic basis for microsatellite instability. *Microsatellites: Evolution and applications*, 1, 34-48.

EKESI, S., NDERITU, P. W., and CHANG, C. L. (2007). Adaptation to and small-scale rearing of invasive fruit fly *Bactrocera invadens* (Diptera: Tephritidae) on artificial diet. *Annals of the Entomological Society of America*, 100(4), 562-567.

ESTOUP, A., GARNERY, L., SOLIGNAC, M., & CORNUET, J. M. (1995). Microsatellite variation in honey bee (*Apis mellifera* L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics*, 140(2), 679-695.

FACHIN, D. A., and DE ASSIS-PUJOL, C. V. (2016). FAMILY STRATIOMYIDAE. *Zootaxa*, 4122(1), 312-341.

FACON, B., HUFBAUER, R. A., TAYEH, A., LOISEAU, A., LOMBAERT, E., VITALIS.*et al.*, (2011). Inbreeding depression is purged in the invasive insect *Harmonia axyridis*. *Current Biology*, 21(5), 424-427.

FAIRCLOTH, B. C. (2008). Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, 8(1), 92-94.

FERGUSON, H. M., JOHN, B., NG'HABI, K., and KNOLS, B. G. (2005). Redressing the sex imbalance in knowledge of vector biology. *Trends in Ecology and Evolution*, 20(4), 202-209.

FERRO, P., R. DELL'EVA, and U. PFEFFER. (2001). Are there CAG repeat expansion-related disorders outside the central nervous system? *Brain Research Bulletin*, 56(3),259–264.

FINNEY, G. L., and FISHER, T. W. (1964). Culture of entomophagous insects and their hosts. In: *DeBach, Paul and Schlinger, E.T. Biological Control of Insect Pests and Weeds*.pp. 328-355. Chapman and Hall, London.

FLINT, J., and MACKAY, T. F. (2009). Genetic architecture of quantitative traits in mice, flies, and humans. *Genome research*, 19(5), 723-733.

FORSTMEIER, W., SEGELBACHER, G., MUELLER, J.C., and KEMPENAERS, B .(2007). Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology*, 16(19), 4039–4050.

FOULLEY, J. L., and OLLIVIER, L. (2006). Estimating allelic richness and its diversity. *Livestock Science*, 101(1), 150-158.

FOX, C. W., K. L. SCHEIBLY, B. P. SMITH, and W. G. WALLIN. (2007). Inbreeding depression in two seed-feeding beetles, *Callosobruchus maculatus* and *Stator limbatus* (Coleoptera: Chrysomelidae: Bruchinae*). Bulletin of Entomological Research*, 97: 49 54.

FOX, C. W., SCHEIBLY, K. L., WALLIN, W. G., HITCHCOCK, L. J., STILLWELL, R. C., and SMITH, B. P. (2006). The genetic architecture of life span and mortality rates: gender and species differences in inbreeding load of two seed-feeding beetles. *Genetics*, 174(2), 763-773.

FRANKHAM, R. (2008). Genetic adaptation to captivity in species conservation programs. *Molecular Ecology*, 17(1), 325-333.

FURMAN, D. P., YOUNG, R. D., and CATTS, P. E. (1959). *Hermetia illucens* (Linnaeus) as a factor in the natural control of *Musca domestica* Linnaeus. *Journal of Economic Entomology*, 52(5), 917-921.

GILCHRIST, A. S., CAMERON, E. C., SVED, J. A., and MEATS, A. W. (2012). Genetic consequences of domestication and mass rearing of pest fruit fly *Bactrocera tryoni* (Diptera: Tephritidae). *Journal of Economic Entomology*, 105(3), 1051-1056.

GLÉMIN, S. (2003). How are deleterious mutations purged? Drift versus non-random mating. *Evolution*, 57, 2678–2687.

HALE, O. M. (1973). Dried *Hermetia illucens* larvae (Diptera: Stratiomyidae) as a feed additive for poultry. *Journal of Georgia Entomology*, 8, 16-20.

HANCOCK, J.M. (1996). Simple sequences and the expanding genome. *Bioessays*, 18(5), 421-425.

HANCOCK, J. M. (1999) Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Microsatellites. Evolution and Applications (eds Goldstein DB, Schlötterer C), pp. 1–9. Oxford University Press, New York.

HECKER, K. H., and ROUX, K. H. (1996). High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. *Biotechniques*, 20(3), 478-485.

HEDRICK, P. (2005). Large variance in reproductive success and the Ne/N ratio. *Evolution*, 59(7), 1596-1599.

HEM, S., TOURE, S., SAGBLA, C., & LEGENDRE, M. (2008). Bioconversion of palm kernel meal for aquaculture: Experiences from the forest region (Republic of Guinea). *African Journal of Biotechnology*, 7(8), 1192-1198

HILL, W. G., and RASBASH, J. (1986). Models of long term artificial selection in finite population. *Genetical Research*, 48(01), 41-50.

HOFFMAN, J. I., & AMOS, W. (2005). Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*, 14(2), 599-612.

HOLMES, L. A., VANLAERHOVEN, S. L., and TOMBERLIN, J. K. (2013). Substrate effects on pupation and adult emergence of *Hermetia illucens* (Diptera: Stratiomyidae). *Environmental Entomology*, 42(2), 370-374.

HOLMES, L. A., VANLAERHOVEN, S. L., and TOMBERLIN, J. K. (2012). Relative humidity effects on the life history of *Hermetia illucens* (Diptera: Stratiomyidae). *Environmental Entomology*, 41(4), 971-978.

HONĚK, A. (1993). Intraspecific variation in body size and fecundity in insects: a general relationship. *Oikos*, 66(3) 483-492.

HOWELL, P. I., and KNOLS, B. G. (2009). Male mating biology. *Malaria Journal*, 8 (2), 1-8.

KALINOWSKI, S. T. (2005). hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, 5(1), 187-189.

KALOVÁ, M., and BORKOVCOVÁ, M. (2013). Voracious larvae *Hermetia illucens* and treatment of selected types of biodegradable waste. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 61(1), 77-83.

KENIS, M., KONÉ, N., CHRYSOSTOME, C. A. A. M., DEVIC, E., KOKO, G. K. D., CLOTTEY, V. A., and MENSAH, G. A. (2014). Insects used for animal feed in West Africa. *Entomologia*, 2(2).

KLINGENBERG, C. P. (2010). Evolution and development of shape: integrating quantitative approaches. *Nature Reviews Genetics*, 11(9): 623–635.

KNIPLING, E.(1955). Possibilities of insect control or eradication through the use of sexually sterile males. *Journal of Economic Entomology*, 48(4), 459–462.

KNOWLES, L. L., LEVY, A., MCNELLIS, J. M., GREENE, K. P., and FUTUYMA, D. J. (1999). Tests of inbreeding effects on host-shift potential in the phytophagous beetle *Ophraella communa*. *Evolution*, 53(2), 561-567.

KRISTENSEN, T. N., SØRENSEN, P., KRUHØFFER, M., PEDERSEN, K. S., and LOESCHCKE, V. (2005). Genome-wide analysis on inbreeding effects on gene expression in *Drosophila melanogaster*. *Genetics*, 171(1), 157-167.

KROECKEL, S., HARJES, A. G., ROTH, I., KATZ, H., WUERTZ, S., SUSENBETH, A., and SCHULZ, C. (2012). When a turbot catches a fly: Evaluation of a pre-pupae meal of the Black Soldier Fly (*Hermetia illucens*) as fish meal substitute—Growth performance and chitin degradation in juvenile turbot (*Psetta maxima*). *Aquaculture*, 364, 345-352.

KURIWADA, T., KUMANO, N., SHIROMOTO, K., and HARAGUCHI, D. (2010). Effect of mass rearing on life history traits and inbreeding depression in the sweetpotato weevil (Coleoptera: Brentidae). *Journal of Economic Entomology*, 103(4), 1144-1148.

LABORDA, P. R., MORI, G. M., and DE SOUZA, A. P. (2009). *Drosophila mediopunctata* microsatellites I: more than a hundred polymorphic loci available for genetic studies. *Conservation Genetics Resources*, 1(1), 297-307.

LAGERCRANTZ, U., ELLEGREN, H., and ANDERSSON, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research*, 21(5), 1111-1115.

LAI, Y., and SUN, F. (2003). The relationship between microsatellite slippage mutation rate and the number of repeat units. *Molecular Biology and Evolution*, 20(12), 2123-2131.

LALANDER, C. H., FIDJELAND, J., DIENER, S., ERIKSSON, S., and VINNERÅS, B. (2015). High waste-to-biomass conversion and efficient Salmonella spp. reduction using black soldier fly for waste recycling. *Agronomy for Sustainable Development*, 35(1), 261-271.

LECLERQ, M. (1969). Disperion et transport des insects nuisbles: *Hermetia illucens*. L. (Diptera Stratiomydiae). *Bulletin des recherches agronomiques de Gembloux*, 4, 139-143.

LI, L., ANDERSON, S., SECOMBE, J., and EISENMAN, R. N. (2013). The *Drosophila* ubiquitin-specific protease Puffyeye regulates dMyc-mediated growth. Development, 140(23), 4776-4787.

LI, Q., ZHENG, L., QIU, N., CAI, H., TOMBERLIN, J.K. and YU, Z. (2011). Bioconversion of dairy manure by black soldier fly (Diptera: Stratiomyidae) for biodiesel and sugar production. Waste Management, 31(6), pp. 1316-1320.

LI, Y. C., KOROL, A. B., FAHIMA, T., and NEVO, E. (2004). Microsatellites within genes: structure, function, and evolution. *Molecular biology and evolution*, 21(6), 991-1007.

LI, Y. C., KOROL, A. B., FAHIMA, T., BEILES, A., and NEVO, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, 11(12), 2453-2465.

LIU, Z. J, and CORDES J.F. (2004). DNA markers and their applications in aquaculture genetics. *Aquaculture*, 238: 1-37.

LORD, W..D., GOFF, M. .L., ADKINS, T.R., HASKELL, N.H. (1994). The black soldier fly *Hermetia illucens* (Diptera: Stratiomyidae) as a potential measure of human postmortem interval: Observations and case histories. *Journal of Forensic Science*, 39(1), 215-222.

LOUKAS, M., ECONOMOPOULOS, A. P., ZOUROS, E., and VERGINI, Y. (1985). Genetic changes in artificially reared colonies of the olive fruit fly (Diptera: Tephritidae). *Annals of the Entomological Society of America*, 78(2), 159-165.

LUNDQVIST, J., DE FRAITURE, C., and MOLDEN, D. (2008). Saving water: from field to fork: curbing losses and wastage in the food chain. Stockholm, Sweden: Stockholm International Water Institute.

MARSHALL, S. A., WOODLEY, N. E., and HAUSER, M. (2015). The historical spread of the Black Soldier Fly, *Hermetia illucens* (L.)(Diptera, Stratiomyidae, Hermetiinae), and its establishment in Canada. *The Journal of the Entomological Society of Ontario*, 146, 51-54.

MEATS, A., HOLMES, H. M., and KELLY, G. L. (2004). Laboratory adaptation of *Bactrocera tryoni* (Diptera: Tephritidae) decreases mating age and increases protein consumption and number of eggs produced per milligram of protein. *Bulletin of entomological research*, 94(06), 517-524.

MESSINA, F. J., MORRIS, B. D., and FOX, C. W. (2013). Effect of inbreeding on host discrimination and other fitness components in a seed beetle. *Annals of the Entomological Society of America*, 106(1), 128-135.

MILLER, M. J., & YUAN, B. Z. (1997). Semiautomated resolution of overlapping stutter patterns in genomic microsatellite analysis. *Analytical Biochemistry*, 251(1), 50-56.

MIYATAKE, T. (1998). Genetic changes of life history and behavioural traits during mass-rearing in the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae). *Researches on Population Ecology*, 40(3), 301-310.

MIYATAKE, T., and YAMAGISHI, M. (1999). Rapid evolution of larval development time during mass-rearing in the melon fly, *Bactrocera cucurbitae*. *Researches on population ecology*, 41(3), 291-297.

NAGARAJU, J. (2002). Application of genetic principles for improving silk production. *Current Science*, 83(4), 409-414.

NAKAGAKI, B. J., and DEFOLIART, G. R. (1991). Comparison of diets for mass-rearing *Acheta domesticus* (Orthoptera, Gryllidae) as a novelty food, and comparison of food conversion efficiency with values reported for livestock. *Journal of Economic Entomology*, 84, 891–896.

NanoDrop® ND-1000 Spectrophotometer User's Manual V3.0.1 (http://microarray.okstate.edu/files/nd-1000-users-manual.pdf)

NEI M. MARUYAMA T. CHAKRABORTY R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* 29: 1-10.

NEWTON, G. L., BOORAM, C. V., BARKER, R. W., and HALE, O. M. (1977). Dried Larvae Meal as a Supplement for Swine. *Journal of Animal Science*, 44(3), 395-400.

NGUYEN, T. T., TOMBERLIN, J. K., and VANLAERHOVEN, S. (2015). Ability of Black Soldier Fly (Diptera: Stratiomyidae) Larvae to Recycle Food Waste. *Environmental entomology*, 44(2), 406-410.

NIJHOUT, H. F., and WILLIAMS, C. M. (1974). Control of moulting and metamorphosis in the tobacco hornworm, *Manduca sexta* (L.): cessation of juvenile hormone secretion as a trigger for pupation. *Journal of Experimental Biology*, 61(2), 493-501.

NORIEGA, F. G. (2014). Juvenile hormone biosynthesis in insects: What is new, what do we know, and what questions remain? *International Scholarly Research Notices*, 2014, 1-16.

NORRIS, D. E., SHURTLEFF, A. C., TOURÉ, Y. T., and LANZARO, G. C. (2001). Microsatellite DNA polymorphism and heterozygosity among field and laboratory populations of *Anopheles gambiae ss* (Diptera: Culicidae). *Journal of Medical Entomology*, 38(2), 336-340.

NOWAK, C., VOGT, C., DIOGO, J. B., and SCHWENK, K. (2007). Genetic impoverishment in laboratory cultures of the test organism *Chironomus riparius*. *Environmental Toxicology and Chemistry*, 26(5), 1018-1022.

NUNNEY, L. (1993). The influence of mating system and overlapping generations on effective population size. *Evolution*, 47, 1329-1341.

OLIVEIRA, F. R., DOELLE, K., and SMITH, R. P. (2016). External Morphology of *Hermetia illucens* Stratiomyidae: Diptera (L. 1758) based on electron microscopy. *Annual Research and Review in Biology*, 9 (5), 1.

OONINCX, D. G. A. B., VOLK, N., DIEHL, J. J. E., VAN LOON, J. J. A., and BELUŠIČ, G. (2016). Photoreceptor spectral sensitivity of the compound eyes of black soldier fly (*Hermetia illucens*) informing the design of LED-based illumination to enhance indoor reproduction. *Journal of Insect Physiology*, 95, 133-139

PARFITT, J., BARTHEL, M., and MACNAUGHTON, S. (2010). Food waste within food supply chains: quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1554), 3065-3081.

PARK S. (2001). The Excel microsatellite toolkit. http://animalgenomics.ucd.ie/sdepark/ms-toolkit

PARKER, A. G. (2005). Mass-rearing for sterile insect release. In: Sterile Insect Technique (pp. 209-232). Springer Netherlands.

PASTOR, B., CICKOVÁ, H., KOZÁNEK, M., MARTÍNEZ-SÁNCHEZ, A., TAKÁC, P., and ROJO, S. (2011). Effect of the size of the pupae, adult diet, oviposition substrate and adult population density on egg production in *Musca domestica* (Diptera: Muscidae). *European Journal of Entomology*, 108(4), 587.

PASTOR, B., MARTÍNEZ-SÁNCHEZ, A. S., STÅHLS, G. A., and ROJO, S. (2014). Introducing improvements in the mass rearing of the housefly: biological, morphometric and genetic characterization of laboratory strains. *Bulletin of entomological research*, 104(04), 486-493.

PEAKALL, R. O. D., and SMOUSE, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular ecology notes*, 6(1), 288-295.

PETERS, M. T., and BARBOSA, P. (1977). Influence of population density on size, fecundity, and evelopmental rate of insects in culture. *Annual Review of Entomology*, 22, 431–450.

PICKER, M., GRIFFITHS, C., and WEAVING, A, 2004 *Field guide to insects of South Africa*. Struik.New Holland Books.

PIRY S., LUIKART G., CORNUET J.M., (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, 90: 502-503.

PRICE, E.O. (1984) Behavioural aspect of animal domestication. *Quarterly review of Biology*, 59, 1-32.

PRIMMER, C. R., RAUDSEPP, T., CHOWDHARY, B. P., MØLLER, A. P., & ELLEGREN, H. (1997). Low frequency of microsatellites in the avian genome. *Genome Research*, 7(5), 471-482.

QUELLER, D. C., and GOODNIGHT, K. F. (1989). Estimating relatedness using genetic markers. *Evolution*, 43(2), 258-275.

RANA, K. J., SIRIWARDENA, S., and HASAN, M. R. (2009). *Impact of rising feed ingredient prices on aquafeeds and aquaculture production* (No. 541). Food and Agriculture Organization of the United Nations (FAO), Rome.

RAUW W.M., KANIS E., NOORDHUIZEN-STASSEN E.N., GROMMERS F.J. (1998). Undesirable side effects of selection for high production efficiency in farm animals: a review. *Livestock Production Science*, 56: 15-33.

RAYMOND M. and ROUSSET F., (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86:248-249.

REED, D. H., & FRANKHAM, R. (2003). Correlation between fitness and genetic diversity. Conservation biology, 17(1), 230-237.

ROHÁČEK, J., and HORA, M, (2013). A northernmost European record of the alien black soldierfly *Hermetia illucens* (Linnaeus, 1758) (Diptera: Stratiomyidae). *Casopis Slezskeho Zemskeho Muzea*, 62(2):101-106.

ROUSSET, F., 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, 8: 103-106.

RUMPOLD, B. A., and SCHLÜTER, O. K. (2013). Potential and challenges of insects as an innovative source for food and feed production. *Innovative Food Science and Emerging Technologies*, 17, 1-11.

SAGHAI-MAROOF, M. A., SOLIMAN, K. M., JORGENSEN, R. A., and ALLARD, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences*, 81(24), 8014-8018.

SANTOS, J., PASCUAL, M., SIMÕES, P., FRAGATA, I., LIMA, M., KELLEN, B., ...and MATOS, M. (2012). From nature to the laboratory: the impact of founder effects on adaptation. *Journal of Evolutionary Biology*, 25(12), 2607-2622.

SCHABEL, H. G. (2010). Forest insects as food: A global review. In P. B. Durst, D. V. Johnson, R. N. Leslie, and K. Shono (Eds.), Forest insects as food: Humans bite back (pp. 37–64). Bangkok, Thailand: FAO.

SCHREMMER, F. (1984). The polymetabol development of the soldier fly larva *Hermetia illucens*.-a contribution to the metamorphosis of the Stratiomyidae. *Annalen des Naturhistorischen Museums in Wien. Serie B für Botanik und Zoologie*, 405-429.

SCRIBER, J. M., and SLANSKY, F. (1981). The nutritional ecology of immature insects. Annual *Review of Entomology*, 26, 183–211.

SELKOE, K. A., and TOONEN, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters*, 9(5), 615-629.

SHARABY, A., MONTASSER, S. A., MAHMOUD, Y. A., and IBRAHIM, S. A. (2010). The possibility of rearing the grasshopper, *Heteracris littoralis* (R.) on semi synthetic diet. Journal of Agriculture and Food Technology, 1, 1–7.

SHEPPARD, D. C., G. L. NEWTON, S. A. THOMPSON and S. SAVAGE. (1994). A value -added manure management-system using the black soldier fly. *Bioresource Technology,* 50, 275-279.

SHEPPARD, D. C., TOMBERLIN, J. K., JOYCE, J. A., KISER, B. C., and SUMNER, S. M. (2002). Rearing methods for the black soldier fly (Diptera: Stratiomyidae).*Journal of Medical Entomology,* 39(4), 695-698.

SHIMOJI, Y., and MIYATAKE, T. (2002). Adaptation to artificial rearing during successive generations in the West Indian sweetpotato weevil, *Euscepes postfasciatus* (Coleoptera: Curculionidae). *Annals of the Entomological Society of America*, 95(6), 735-739.

SIMÕES, P., SANTOS, J., FRAGATA, I., MUELLER, L. D., ROSE, M. R., and MATOS, M. (2008). How repeatable is adaptive evolution? The role of geographical origin and founder effects in laboratory adaptation. *Evolution*, 62(8), 1817-1829.

SINGH, P. (1982). The rearing of beneficial insects. *New Zealand Entomologist*, 7, 304–310.

SMYKAL, V., DAIMON, T., KAYUKAWA, T., TAKAKI, K., SHINODA, T., and JINDRA, M. (2014). Importance of juvenile hormone signaling arises with competence of insect larvae to metamorphose. *Developmental biology*, 390(2), 221-230.

SONG, S. D., DREW, R. A., and HUGHES, J. M. (2007). Multiple paternity in a natural population of a wild tobacco fly, *Bactrocera cacuminata* (Diptera: Tephritidae), assessed by microsatellite DNA markers. *Molecular Ecology*, 16(11), 2353-2361.

SØRENSEN, J. G., ADDISON, M. F., and TERBLANCHE, J. S. (2012). Mass-rearing of insects for pest management: Challenges, synergies and advances from evolutionary physiology. *Crop Protection,* 38, 87-94.

SOUZA, H. M. L., MATIOLI, S. R., and SOUZA, W. N. (1988). The adaptation process of *Ceratitis capitata* to the laboratory analysis of life-history traits. *Entomologia experimentalis et applicata*, 49(3), 195-201.

ST-HILAIRE, S., CRANFILL, K., MCGUIRE, M. A., MOSLEY, E. E., TOMBERLIN, J. K., NEWTON, L., ... and IRVING, S. (2007). Fish Offal Recycling by the Black Soldier Fly Produces a Foodstuff High in Omega-3 Fatty Acids. *Journal of the World Aquaculture Society*, 38(2), 309-313.

TCHUINKAM, T., MPOAME, M., MAKE-MVEINHYA, B., SIMARD, F., LELE-DEFO, E., ZEBAZE-TOGOUET, S., *et al*. (2011). Optimization of breeding output for larval stage of *Anopheles gambiae* (Diptera: Culicidae): Prospects for the creation and maintenance of laboratory colony from wild isolates. *Bulletin of Entomological Research*, 101, 259–269.

THOREN P.A., PAXTON R.J., ESTOUP A. (1995). Unusually high frequency of (CT)n and (GT)n microsatellite loci in a yellowjacket wasp, *Vespula rufa* (L.) (Hymenoptera: (Hymenoptera: Vespidae). *Insect Molecular Biology*, 4, 141-148.

TOMBERLIN, J. K., and SHEPPARD, D. C. (2001). Lekking behaviour of the black soldier fly (Diptera: Stratiomyidae). *Florida Entomologist*, 84(4), 729-730.

TOMBERLIN, J. K., and SHEPPARD, D. C. (2002). Factors influencing mating and oviposition of black soldier flies (Diptera: Stratiomyidae) in a colony. *Journal of Entomological Science*, 37(4), 345-352

TOMBERLIN, J. K., SHEPPARD, D. C., and JOYCE, J. A. (2002). Selected life-history traits of black soldier flies (Diptera: Stratiomyidae) reared on three artificial diets. *Annals of the Entomological Society of America*, 95(3), 379-386.

TRECO, D. and ARNHEIM, N. (1986). The evolutionarily conserved repetitive sequence d(TG. AC)n promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. *Molecular and Cellular. Biology*. 6, 3934–3947.

UNTERGASSER, A., NIJVEEN, H., RAO, X., BISSELING, T., GEURTS, R., and LEUNISSEN, J. A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic acids research*, 35(suppl 2), W71-W74.

ÜSTÜNER, T., HASBENLÍ, A., and ROZKOŠNÝ, R. (2003). The first record of *Hermetia illucens* (Linnaeus, 1758) (Diptera, Stratiomyidae) from the Near East. *Studia dipterologica*, 10(1), 181-185.

VAN HUIS, A. (2013). Potential of insects as food and feed in assuring food security. *Annual Review of Entomology*, 58, 563-583.

VAN OOSTERHOUT C,. HUTCHINSON W.F., WILLS D.P.M., SHIPLEY P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4: 535-538.

VAN OOSTERHOUT. C., WEETMAN. D.. and HUTCHINSON. W. F. (2006). Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes*. 6(1). 255-256.

VANTOMME, P., MERTENS, E., VAN HUIS, A., and KLUNDER, H. C. (2012). Summary report of the technical consultation meeting "Assessing the Potential of Insects as Food and Feed in assuring Food Security. In FAO (Ed.). Rome.

VICOSO, B., and BACHTROG, D. (2015). Numerous transitions of sex chromosomes in Diptera. *PLoS Biol*, 13(4), e1002078.

VILLET, M. H. (2011). African carrion ecosystems and their insect communities in relation to forensic entomology. *Pest Technology*, 5(1), 1-15.

WAITS, L. P., LUIKART, G., and TABERLET, P. (2001). Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, 10(1), 249-256.

WANG, J. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics*, 166(4), 1963-1979.

WANG, M. L., BARKLEY, N. A., & JENKINS, T. M. (2009). Microsatellite markers in plants and insects. Part I: Applications of biotechnology. *Genes, Genomes and Genomics*, 3(1), 54-67.

WANG, Q., and WANG, X. (2012). Comparison of methods for DNA extraction from a single chironomid for PCR analysis. *Pakistan Journal of Zoology*, 44, 421-426.

WAPLES, R. S., and DO, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3(3), 244-262.

WEIR, B. S., & COCKERHAM, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358-1370.

WELDON, C. W. (2005). Mass-rearing and sterilisation alter mating behaviour of male Queensland fruit fly, *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae). *Australian Journal of Entomology*, 44(2), 158-163.

WILLOUGHBY, J. R., FERNANDEZ, N. B., LAMB, M. C., IVY, J. A., LACY, R. C., and DEWOODY, J. A. (2015). The impacts of inbreeding, drift and selection on genetic diversity in captive breeding populations. *Molecular Ecology*, 24(1), 98-110.

WOODLEY, N.E. (2009). Stratiomyidae. In: Brown B.V., Borkent, A., Cumming, J.M., Wood, D.M., Woodley, N.E. & Zumbado, M.A. (Eds.), Manual of Central American Diptera. Vol. 1. NCR Research Press, Boca Raton, Ottawa, pp. 521–549.

WOODS, J. G., PAETKAU, D., LEWIS, D., MCLELLAN, B. N., PROCTOR, M., and STROBECK, C. (1999). Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin*, 616-627.

ZANE, L., BARGELLONI, L., and PATARNELLO, T. (2002). Strategies for microsatellite isolation: a review. *Molecular Ecology*, 11(1), 1-16.

ZEH, D. W., ZEH, J. A., and BERMINGHAM, E. (1997). Polyandrous, sperm-storing females: carriers of male genotypes through episodes of adverse selection. *Proceedings of the Royal Society of London B: Biological Sciences*, 264(1378), 119-125.

ZHENG, L., HOU, Y., LI, W., YANG, S., LI, Q., & YU, Z. (2012). Biodiesel production from rice straw and restaurant waste employing black soldier fly assisted by microbes. *Energy*, 47(1), 225-229.

ZHOU, F., TOMBERLIN, J. K., ZHENG, L., YU, Z., and ZHANG, J. (2013). Developmental and waste reduction plasticity of three black soldier fly strains (Diptera: Stratiomyidae) raised on different livestock manures. *Journal of Medical Entomology*, 50(6), 1224-1230.

ZOUROS, E., and LOUKAS, M. (1989). Biochemical and colonization genetics of *Dacus oleae* (Gmelin). *Fruit flies: their biology, natural enemies and control*, 3, 75-87.

ZYGOURIDIS, N. E., ARGOV, Y., NEMNY-LAVY, E., AUGUSTINOS, A. A., NESTEL, D., and MATHIOPOULOS, K. D. (2014). Genetic changes during laboratory domestication of an olive fly SIT strain. *Journal of applied entomology*, 138(6), 423- 432.

Appendix A.

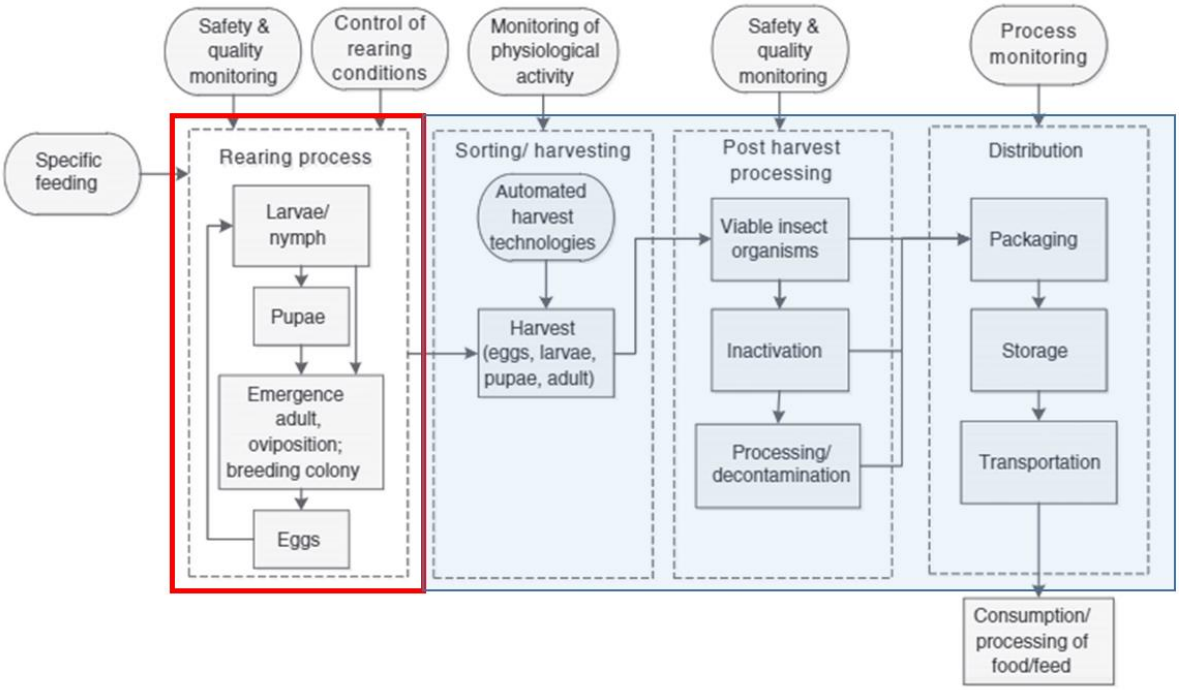

Figure A.1: Production process of food and feed derived from edible insects. Highlighted in red is the rearing process, applicable to most insect mass-rearing facilities (Rumpold and Schlüter, 2013)

Appendix B.

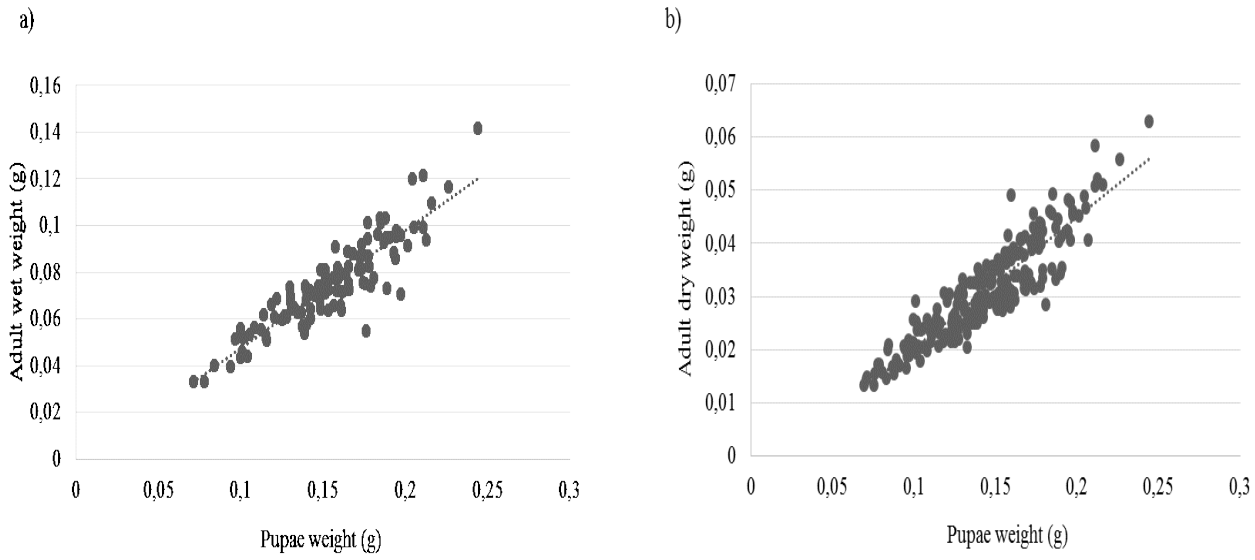

Figure B.1: Two hundred 6d/o larvae were fed layer hen feed (Quantum Foods (Pty) Ltd t/a Nova Feeds) at different feeding rates (50, 100 and 200mg/day) for 12 days in 1L plastic tubs in order to achieve a wide spectrum of individual sizes. Pearson's correlations illustrate the relationship between pupae weight and wet (a) (r = 0.8989; R^2=0.8080; P = <0.01) and dry (b) adult weight (r = 0.9499; R^2=0.8274; P < 0.01).

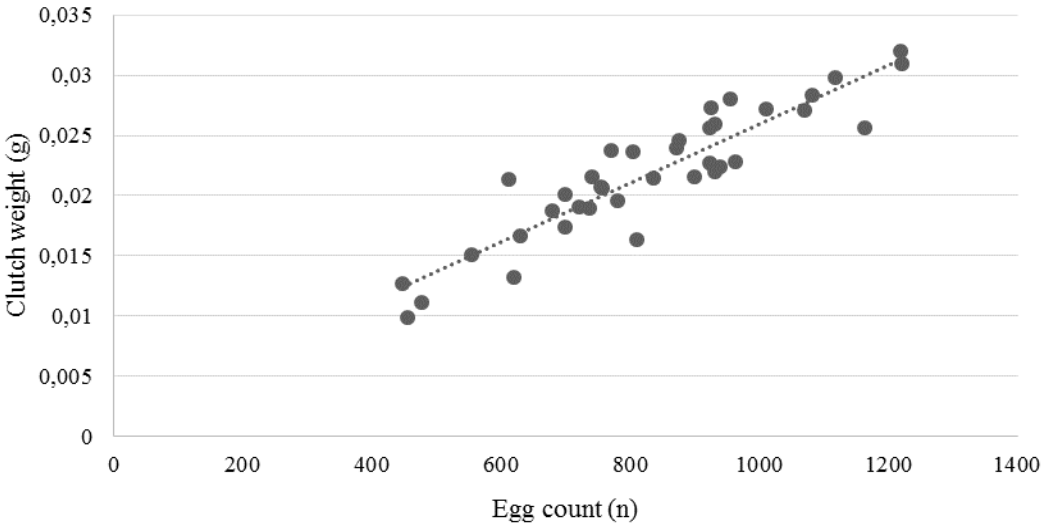

Figure B.2: Pearson's correlation illustrate the correlation between number of eggs in a clutch and the total weight of each clutch (n = 38; r =-0.9433; R^2 = 0.8372; P = < 0.05). The average size of a single clutch of eggs laid by a female was determined by sampling 50 egg clutches.

Table B.1: compilation of data from Holmes *et al.* (2012) and the current study, illustrating the relationship between hatchability and relative humidity.

Relative Humidity (%)	Average Hatchability (%)
25[a]	8.44±0.75
25[a]	5.44±1.4
70[a]	64.67±7.69
70[a]	86.22±2.91
40[a]	19.86±1.27
50[a]	38±2.29
60[a]	72.74±2.46
65.8[b]	48±36.8
70.8[b]	91±7.1
69[b]	68±30.88
63[b]	60±21.00

[a]Data from Holmes *et al.*, (2012)

[b]Data from Chapter 2

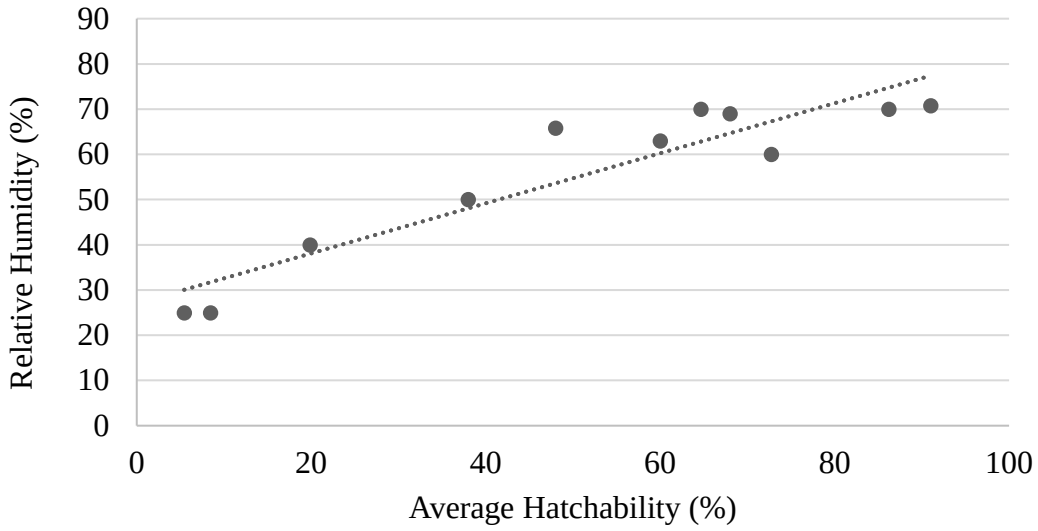

Figure B.3: A significant correlation exists between relative humidity (% RH) and egg hatchability. Data from a previous publication was plotted in combination with data from this study by implementing Pearson's correlations (r = 0.9300; R^2 = 0.866 $P \leq 0.001$)

Appendix C.

Table C.1: *De novo* genome assembly statistics as carried out on CLC Genomic Workbench (v7.0.4) for *Hermetia illucens*.

Assembly Statistics	Male Length (bp)	Female Length (bp)	Combined Length (bp)
N50 (bp)	750	747	838
Minimum (bp)	500	500	500
Maximum (bp)	15,698	15,698	6,977
Average (bp)	769	767	862
Count	50,419	46,836	108,770
Total (bp)	38,749,532	35,944,941	93,759,357

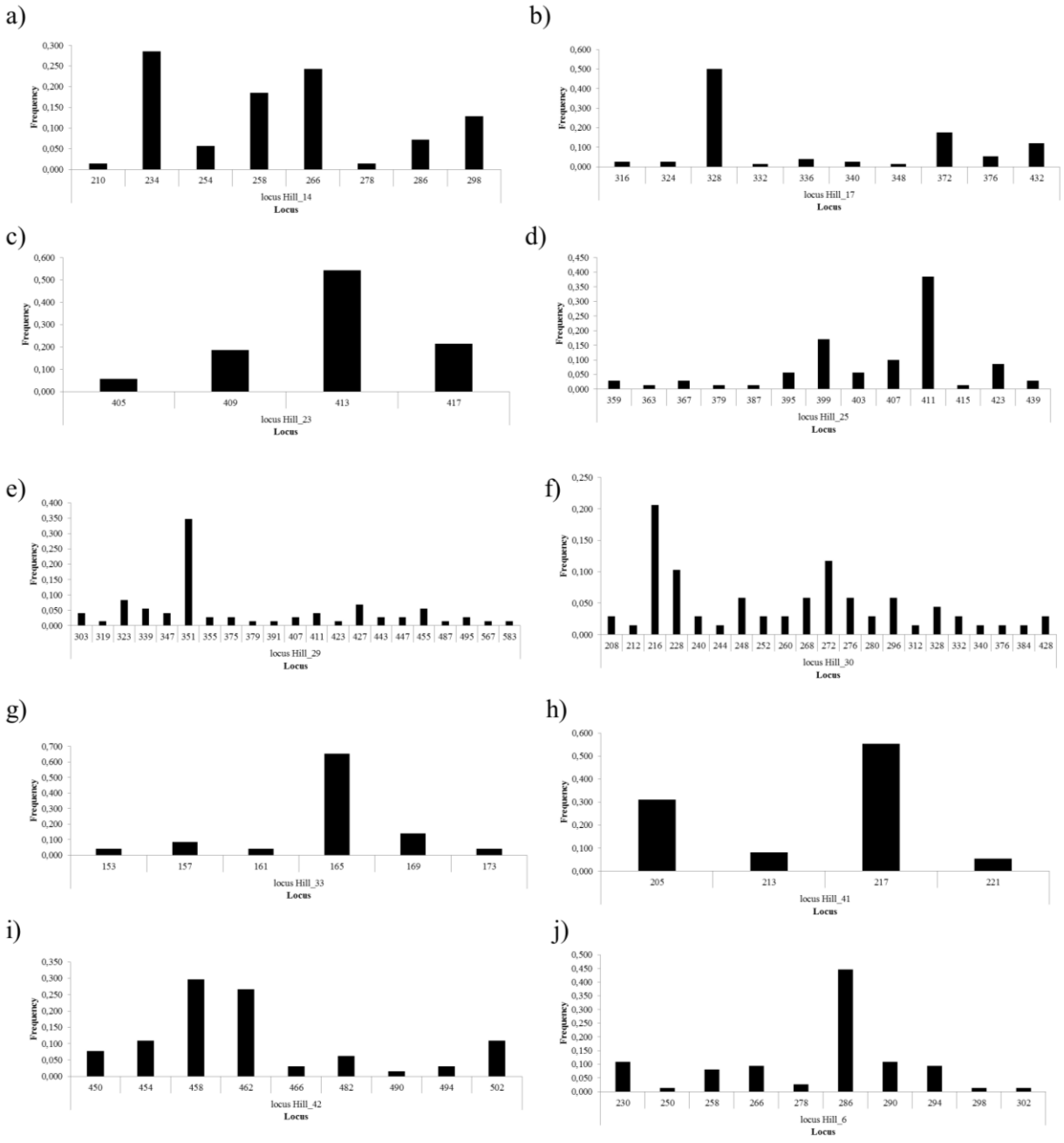

Figure C.1: Allele frequency distributions of ten polymorphic microsatellite loci for *Hermetia illucens*. Presented are a) Hill_14, b) Hill_17, c) Hill_23, d) Hill_25, e) Hill_29, f) Hill_30, g) Hill_33; h) Hill_41, i) Hill_42 and j) Hill_6.

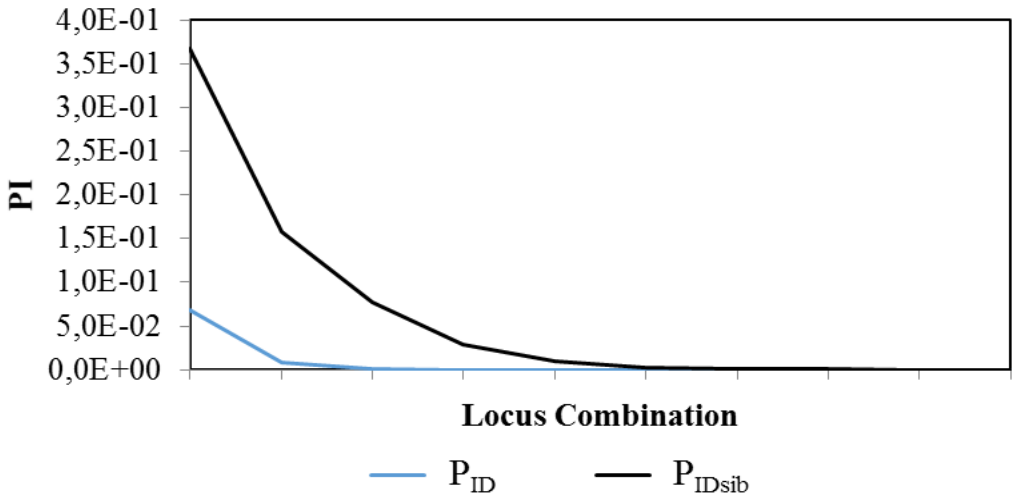

Figure C.2: The probability of identity in the sampled *Hermetia illucens* population for increasing combination of ten polymorphic microsatellite loci. (P_{ID}) probability of identity of two independent sample and amongst siblings (P_{IDsib}), as computed on GenAlex v.6.41 (Peakall and Smouse, 2006).

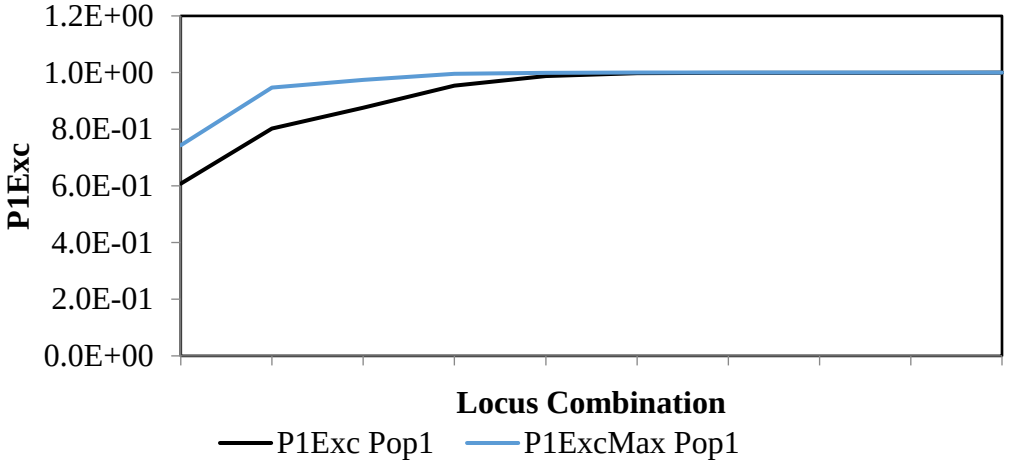

Figure C.3: Exclusion probabilities in the sampled *Hermetia illucens* population for increasing combinations of ten polymorphic microsatellite loci when the genotype of one parent is known as computed on GenAlex v.6.41 (Peakall and Smouse, 2006).

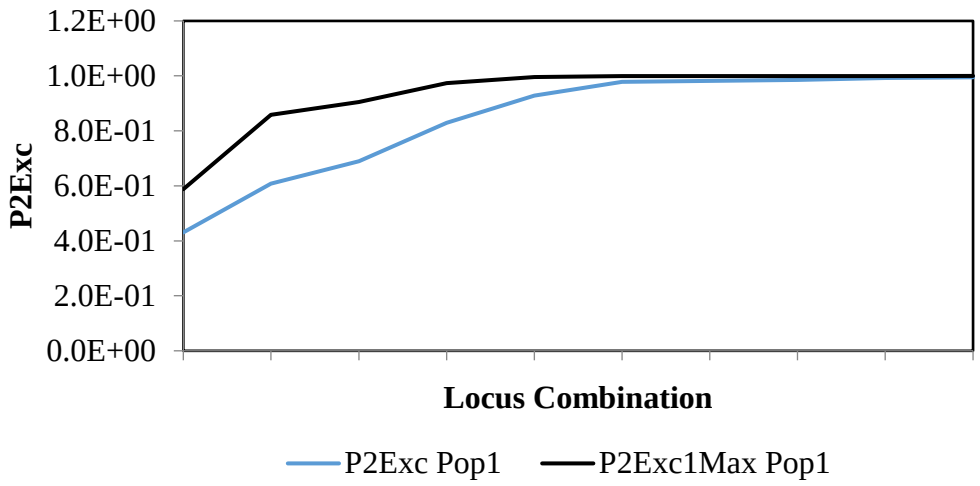

Figure C.4: Exclusion probabilities in the sampled *Hermetia illucens* population for increasing combinations of ten polymorphic microsatellite loci when the genotype of both parents is known as computed on GenAlex v.6.41 (Peakall and Smouse, 2006).

Appendix D.

Table D.1: Additional genetic diversity statistics for six experimental (F0-F5) and one commercial (F28) cohort of *Hermetia illucens*. N= sample size; N_A = number of alleles; A_E= number of effective alleles; I = Shannon's diversity index; H_E = expected heterozygosity; H_O = observed heterozygosity; uH_E = unbiased expected heterozygosity; R_S = allelic richness; R_{SP} = private allelic richness, F_{IS} = fixation index, F_{NULL} = null allele frequencies as determined by the Brookfield 2 method (Brookfield, 1996)

	N	N_A	A_E	I	H_E	H_O	uH_E	R_S	R_{SP}	F_{IS}	F_{NULL}
F0	35	10.600	4.737	1.751	0.736	0,477	0,747	2,84	1,03	0.4772	0.3666
F1	30	6.600	2.547	1.125	0.544	0,509	0,554	2,22	0,33	0.5091	0.2336
F2	24	5.600	2.796	1.061	0.528	0,437	0,539	2,18	0,28	0.4367	0.5525
F3	29	5.700	2.260	0.914	0.431	0,315	0,439	1,97	0,24	0.3150	0.2279
F4	27	5.300	2.277	0.845	0.399	0,343	0,407	1,9	0,21	0.3428	0.04688
F5	19	6.400	3.168	1.201	0.546	0,311	0,559	2,34	0,73	0.3105	0.8477
F28	28	6.000	2.794	1.078	0.506	0,213	0,515	2,18	0,64	0.2129	0.6404

Table D.2: Mean within-population pairwise relatedness estimates as per Queller and Goodnight (1989) as represented by figure D.1, significance values are indicated by *P* at a confidence interval of 95%. Upper (U) and lower (L) confidence intervals are also presented.

	F0	F1	F2	F3	F4	F5
Mean	-0.101	0.178	0.128	0.289	0.340	0.070
U	0.061	0.063	0.082	0.063	0.077	0.068
L	-0.091	-0.098	-0.110	-0.099	-0.108	-0.104
P	0.992	0.001	0.007	0.001	0.001	0.024

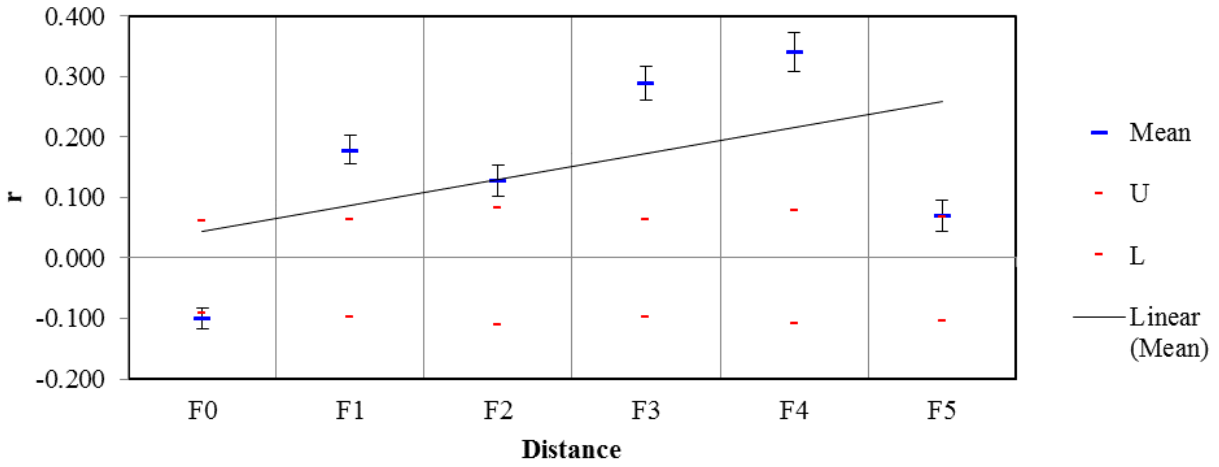

Figure D.1: Mean within-population pairwise relatedness estimates computed as per Queller and Goodnight (1989). Mean relatedness is demarcated in blue ($y = 0.0428x+0.0008$; $R^2 = 0.2546$); red demarcations indicate upper- and lower confidence intervals (95%). The generational cohorts are represented as F0 to F5.

Table D.3: Pearson (*r*) and Spearman's Rho (ρ) correlation coefficients are listed below for various mean phenotypic traits and mean measures of genetic diversity for each experimental generation cohort, **R_S = allelic richness; F_{IS} = Wright's F statistic; N_A = number of alleles; H_O = observed heterozygosity**. Significance values are indicated by "*P*".

		Pearson's *r*		Spearman's ρ	
		r	*P*	ρ	*P*
R_s	Hatchability	0,142	0,858	0,04	0,92
Relatedness	Hatchability	-0,442	0,558	0,16	0,75
F_{IS}	Hatchability	0,415	0,585	0,64	0,333
N_A	Hatchability	-0,521	0,479	0,16	0,75
H_O	Hatchability	-0,137	0,863	0,16	0,75
R_s	Clutch size	-0,515	0,374	0,64	0,133
Relatedness	Clutch size	0,478	0,415	0,36	0,35
F_{IS}	Clutch size	0,864	0,582	0,09	0,683
N_A	Clutch size	-0,458	0,438	0,36	0,35
H_O	Clutch size	-0,912	0,031*	1	0,017*
R_s	Pupae weight	-0,895	0,040*	0,49	0,233
Relatedness	Pupae weight	0,831	0,082	0,01	0,95
F_{IS}	Pupae weight	0,98	0,535	0	1
N_A	Pupae weight	-0,904	0,035*	0,16	0,517
H_O	Pupae weight	-0,665	0,221	0,64	0,333
R_s	Pupation	-0,586	0,299	0,761	0,083
Relatedness	Pupation	0,613	0,613	0,445	0,233
F_{IS}	Pupation	0,776	0,654	0,042	0,783
N_A	Pupation	-0,453	0,443	0,516	0,133
H_O	Pupation	-0,98	0,003**	0,95	0,017*
R_s	Oviposition	0,668	0,218	0,81	0,083
Relatedness	Oviposition	-0,725	0,166	0,49	0,233
F_{IS}	Oviposition	-0,539	0,925	0,25	0,45
N_A	Oviposition	0,516	0,373	0,64	0,133
H_O	Oviposition	0,925	0,024*	0,81	0,083

"*" indicates significance to the 5% nominal level
"**" indicates significance to the 1% nominal level